VOL. XXXIV NO. 7 — PSYCHOLOGICAL REVIEW PUBLICATIONS — WHOLE NO. 160 1925

Psychological Monographs

The Influence of the Factor of Intelligence on the Form of the Learning Curve

By

GILES MURREL RUCH

University of Iowa

PUBLISHED FOR
The American Psychological Association
By THE PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.
and ALBANY, N. Y.

Agents: G. E. STECHERT & CO., London (2 Star Yard, Carey St., W. C.)
Leipzig (Hospital St., 10); Paris (76, rue de Rennes)

ACKNOWLEDGMENTS

In the course of this investigation the writer has necessarily become deeply indebted to a number of persons. It is a great pleasure to make specific acknowledgment of the invaluable assistance given by Professors Lewis M. Terman and Truman L. Kelley of Stanford University. The completion of the study was made possible through the granting of a University Fellowship for the year 1920–21 and a Research Fellowship for the year 1921–22. Acknowledgment is due the Committee on Graduate Study and to Professor Terman for these awards. Professor William T. Root, Jr., now of the University of Pittsburgh, gave much appreciated aid in the formulation of the experimental procedure. Through the coöperation of Dr. Virgil E. Dickson, Mr. William M. Greenwell, Mr. Charles E. Keyes, and Mr. John R. Sutton, a large number of children in the city schools of Oakland, California, were made available for subjects in the experiments. Mr. Cecil R. Brolyer, of Stanford University, gave valuable assistance in supervising the compilation and treatment of the experimental results. Dr. Arthur S. Otis, of the World Book Company, was kind enough to read and criticize certain parts of the statistical treatment.

CONTENTS

Chap.		Page
I.	Introduction and Statement of the Problem...	1
II.	The Literature of the Problem...............	4
III.	Description of the Methods of the Investigation....................................	14
IV.	The Experimental Results..................	26
V.	Statistical Treatment and Interpretation of the Results..............................	27
VI.	Summary and Conclusions....................	57
	Appendix....................................	59
	Bibliography................................	62

CHAPTER I

Introduction and Statement of the Problem

The present investigation is primarily an attempt to bring under control and measurement certain factors which are operative in producing individual differences during the practice of mental functions. Most experimental studies which have taken cognizance of such individual variations at all have done so in one or both of two principal ways. The first of these problems is that of the quantitative measurement of the amounts of individual variability in the ability to profit by training, and the second is the question whether practice *increases* or *decreases* these differences in the performance of various subjects.

With respect to the first of these two issues it is now common knowledge that relatively great differences in such capacities do exist and the present interest in this question has narrowed down to the statistical expression of these differences in terms of valid objective units. The second question, on the contrary, has not yielded an unambiguous answer and no generalizations of great significance have been forthcoming. Many psychologists have held firmly that practice decreases or levels down those differences, innate or learned, which are present in the initial performances in the formation of new sets of nerve connections. To-day, the opinion tends to the position that these differences are increased by further exercise of the functions involved.

The experimentation which is reported here was planned to attack this problem of the increase or decrease of individual differences under practice in such a way as to control and measure the influence of two factors which are believed to be prominently involved in the behavior and fate of these differences. Specifically stated, the problem is to study: (1) the influence of the type of mental function which is involved in the learning, and (2) the influence of the factor of general intelligence or the general mental level of the subjects used.

In the main, learning studies have involved very small numbers of subjects who were usually adults. For this reason, probably, the question of the possible influence of innate or acquired differences has not ordinarily been raised. There are surprisingly few studies of the learning of children and almost no investigations where the subjects have been carefully classified and measured with reference to the general mental ability of the learners. Conceivably this might prove to be a factor of great significance. As a matter of fact teachers have seemed more aware of the importance of intelligence in learning than have the laboratory investigators.

In the second place there is probably an unwarranted tendency to generalize too largely about *the learning curve* as if *one particular form of a curve* obtained for all types of mental functions. We have, indeed, suggested several general categories of learning types, such as "sensorimotor," "perceptual," "ideational," etc., without raising the question of the possibility of these being found to possess widely divergent characteristics with respect to the form of their learning curves. It is well within the bounds of reason to expect that the intellectual superiority of a gifted individual might not give him as great an advantage over a less gifted individual in his performance in such a relatively simple situation as cancellation of a's or sorting cards as in the mental multiplication of three-place numbers or in a completion exercise. If this be true, it is of little gain to generalize about the course of individual differences under practice without specific reference to the kind of learning involved.

To measure the influence of these two factors, viz., general mental ability and the type of mental function undergoing practice, the selection of the subjects in the present experiment had to be carried out in a definite manner. The chronological age range was kept as small as was practicable under the actual conditions of securing the necessary subjects. At the same time the mental age range was made as great as conveniently possible in order that the extremes of mental ability might be well represented in numbers.

The selection of suitable tests offered considerable difficulty. In order to vary the type of mental processes to be studied, in accordance with the previously discussed requirements, and not make the experimental work prohibitive in amount, it was desirable to select a small number of tests which would, in a sense, sample the whole range of possibilities from " motor " learning on the one hand to the so-called " higher thought processes " on the other. Tests of sensorimotor capacities are fairly numerous, and one, card sorting, which has been used by many investigators (Bergström, Burt, Coover, Brown, *et al.*) was finally selected as representative of its type. As a second test which would involve comparatively little of the motor element and which would call for use of sensational and perceptual processes, a modification of the Healy Civil War Code was adopted. The selection of a third test needed for the study of learning on a highly intellectual level proved to be very difficult. The possibilities included such tests as the analogies, completion exercises, mental multiplication or addition, and a few others. Unfortunately none of these exist in several comparable forms and it is difficult to measure the efficiency of the learning of such materials from day to day. The final choice was that of an abstract mathematical relations test which will be described in greater detail in a later section. No claim is made that the location of these three tests with respect to their psychological characteristics and demands is based upon any more exact knowledge than a considerable amount of traditional agreement in the terminology of the literature of learning. The real intention, as has already been stated, was merely that of sampling a fairly wide range of mental capacities.

CHAPTER II

The Literature of the Problem

The attempt to summarize even the important experimental work that has been done on the general problem of individual differences in learning capacities would not be warranted by the pertinence of much of this literature to present needs. Nevertheless reference must be made to a number of the older studies which have only indirect bearing in order to give the proper historical orientation for certain controversies which must be taken up later. Recently there has grown up something of a literature on the relation of intelligence to learning. This demands somewhat more careful consideration.

For these reasons, the selection of the papers to be reviewed will be restricted to three groups, viz.: (1) Representative older studies of the increase or decrease of individual differences during practice, of the relations existing between initial and final efficiency in learning, and of the course of the behavior of intercorrelations between mental functions under practice; (2) studies of card sorting, substitution, and reasoning abilities in tests similar to those of the present study; and (3) studies of the learning of subjects classified according to general mental ability, or where the learning has been measured with statistical reference to general intelligence.

Obviously only the investigations of the third type are direct in their significance for present purposes. References to the literature will be made in accordance with the foregoing classification.

I. *Investigations of the Effects of Practice on Individual Differences:*

Binet (1899:4) concluded that for several forms of cancellation tests the differences between bright and dull pupils present at the outset tended to disappear with continued practice. He

says: "La différenciation est nette surtout dès la première épreuve; elle diminue et peut même s'effacer aux épreuves subséquentes" (p. 395).

This conclusion was attacked by Spearman and Krüger (1907: 39) on the basis of their results obtained by refiguring data gathered by Oehrn in 1899 on continued adding. They replied to Binet as follows: "Binet hat, wie schon gesagt, die Meinung ausgesprochen, dass derartige Korrelationen nur bei ungewohnten Versuchsbedingungen deutlich hervortreten; mit zunehmender Uebung sollen sie rasch kleiner werden, manchmal sogar verschwinden. Der eine von uns wurde jedoch zu dem entgegengesetzten Schlusse geführt, dass bei genauer Untersuchtsmethode die Uebung . . . die Korrelationen sogar vergrössert" (p. 96).

Burt (1909:7) studied the same problem but discussion of his results, which are not very conclusive, will be reserved until a later section.

From 1908 on, Thorndike and his students have carried out a series of investigations of individual differences in mental capacities. Thorndike's own paper (1908:46) presented the general tenor of all of the conclusions reached in this group of studies when he says, with reference to multiplication of three-place numbers, ". . . the larger individual differences increase with equal training, showing a positive correlation of high initial ability with ability to profit by training" (p. 384). During the five years from 1911 to 1916, Donovan and Thorndike (1913:15), Hahn and Thorndike (1914:18), Starch (1911:41), Kirby (1913:24), and Thorndike himself in several studies (1910:49, 1914:47, 1915:48), brought forward additional evidence that practice in mental operations in arithmetic resulted in increasing the differences present at the outset. In one of these, (1914:47) after reviewing the work of Galton, Cattell, Rice, and his own study of men entering learned professions, he summarizes by saying: "The facts are rather startling. Equalizing practice *seems to increase* differences. The superior man seems to have got his present superiority by his own nature rather than by

superior advantages of the past, since, during a period of equal advantages for all, he increases his lead" (p. 305).

Wells (1912:52) reported that the learning curves of ten different subjects rarely crossed, and that ". . . a superior performance at the beginning of special practice is not necessarily or even probably attained at the sacrifice of prospects for future improvement" (p. 88).

Hollingworth, (1913:19) from the use of adding, color naming, opposites, discrimination reaction to colors, coördination (3-hole test), and tapping, concluded that the average intercorrelations of the six tests increased from trial to trial, at least up to certain stages. This point was about the twenty-fifth trial for discrimination and the eightieth trial for adding. The average intercorrelations were .065, .280, .320, .390, and .490 for the medians of trial 1, trials 1–5, 20–25, 75–80, and 200–205, respectively. Hollingworth warns against the use of the initial trials as measures of ability and favors trials nearer the limit of efficiency. Cancellation showed a correlation of .665 between initial and final efficiency but opposites yielded but —.088 between the preliminary and 130th trials. The coefficient for adding was .154 under the same conditions. The number of subjects was but thirteen and all were adults. Hollingworth himself has pointed out that cognizance must be taken of the possible effects of changes in the test functions due to habituation during the long practice periods in interpreting his results.

Whitely (1911:54) reported the results of nine adults of varying levels of ability in tests of discrimination of weights, cancellation of A's, sorting, and the pencil maze. Correlations of about .50 between starting ability and gross gain were computed. These results are not in accord with the findings of others and even certain of Whitely's other findings and have been objected to by Thorndike and others.

Chapman (1914:9) found high correlations between initial and final performances in color naming, cancellation, opposites, and multiplication, the coefficients ranging from .59 to .96 in a group of twenty-two male college students.

Jones (1917:23) has held that the results of Wells, Chapman, and Hollingworth are opposed to Thorndike's position. He refers specifically (p. 20) to the fact that the Wells subjects who gained most in addition gained least in cancellation. This raises an issue which will be returned to after our new data have been presented since it is typical of one of the controversial interpretations of experimental results with which the literature of learning abounds.

Strickland, (1918:43) using the Woodworth-Wells color-naming test, adding, tapping, multiplying, and word building, found that "where practice improves performances correlations increase". (p. 399).

A belief in a "general capacity for learning" has been expressed by Pyle (1919:35). Such a capacity is dependent upon the characteristics of the nervous system itself and would operate to cause intercorrelations between separate tests to approach unity if the extraneous factors could be eliminated.

Many apparent inconsistencies are to be noted in the foregoing account, and the conclusions are by no means unambiguous. The variety of mental processes measured, the probable great differences of mental ability among the subjects used, and the variations in the points of view and statistical methods have all contributed to produce these conflicting interpretations. On the whole there seems to be a preponderance of evidence in favor of the view that practice tends to increase those differences present in human beings at the beginning of learning situations, at least for such complex mental functions as the mental solution of arithmetical problems and reasoning abilities in general. For tasks like cancellation and sensory discrimination, the evidence is uncertain and the formulation of definite conclusions is better omitted until our new data have been presented.

II. *Investigations of Card Sorting, Substitution, and Reasoning Abilities:*

Card sorting as a test of mental ability has been used by many investigators for one purpose or another, *e.g.*, by Bergström

(1894:3), Coover and Angell (1907:11), Culler (1912:12), Calfee (1913:8), Kline and Owens (1913:25), Brown (1914:6), Myers (1918:30), Pyle (1919:35), *et al.* Bergström, Culler, Kline and Owens, and Brown were chiefly interested in studying the effects of interference in habit formation. Coover and Angell emphasized the great individual differences in the imagery and other mental processes of the several subjects and in the same subject at different stages in the learning process. However, none of these studies involved card-sorting techniques which resembled at all closely the one used in the present study.

Calfee has reported correlations between the sorting of cards face up and the sorting the same face down (dealing) ranging from .45 to .71. Of card sorting, card dealing, alphabet sorting, and mirror writing, the first mentioned gave the highest correlation with school grades of children.

Myers has stated that "practice does not make the individuals more or less alike" (p. 325). He refers to card sorting.

Substitution tests have been used in a variety of forms by many workers, *e.g.*, Gray (1918:17), Baldwin (1913:2), Starch (1912:42), Lough (1912:27), Munn (1909:29), Pyle (1913:34), Squire (1912:40), Woodworth and Wells (1911:57), Woolley and Fischer (1914:58), Dearborn and Brewer (1918:14), and many others. The last mentioned only used the Healy Civil War Code. Their methods differed so greatly from those of this investigation that comparisons cannot be made even in this case. Dearborn and Brewer's work, although chiefly intended as a university class demonstration, brought forth certain results of interest, viz., ". . . the students tend to hold the same relative rank in the first trials as in the last trials of the practice" (p. 81). Whipple (1915:53) gives a good summary of the literature and norms on substitution tests (pp. 499–515). Weidensall (1916:51) used the Woolley and Fischer substitution test with delinquent women and found a correlation of .48 with estimated intelligence after several practices.

Bonser (1910:5), Ruger (1910:37), Peterson (1920:33), and many others, have worked with tests of reasoning capacities.

None of these resembles at all closely the "Abstract Relations" test of our experiments and hence need not be commented on at length. Bonser presents evidence that the abilities of naming opposites, controlled association, selective judgments, and even literary interpretation are closely related in their psychological nature.

Ruger has carried out an introspective analysis of the puzzle-solving consciousness with trained adult observers.

III. *Studies of Learning in Relation to General Intelligence.*

Dating from the time of Binet there have been perhaps somewhat more than a score of studies which directly or indirectly have involved a certain amount of consideration of the influence of intelligence as a factor in learning. Binet has already been quoted as believing that dull subjects in certain tests (cancellation) reach a final efficiency equal to that of bright ones.

Kuhlmann (1904:26) found that the learning curves of three Mongolian imbeciles and six feebleminded in target throwing and tapping maintained their relative ranks rather closely.

Terman's study (1906:45) of seven "bright" and seven "stupid" boys presents a detailed account of the differences in the capacities of these boys in the higher mental processes, *e.g.,* in puzzle solving, tests of invention, the ball and field test, mathematical problems, language usage, mutilated texts, fables, memory, chess playing, *et al.* The "bright" group proved much superior to the "stupid" in all the mental tests used, there being the least difference in the tests of invention. The subjects maintained their relative ranks with great uniformity.

Burt (1909:7), with a battery of twelve tests (sensory discrimination, tapping, card sorting, mirror writing, spot-pattern, dotting, etc.), tested English school children in considerable numbers and found that eleven of the twelve tests gave a lower correlation with imputed intelligence on the second trial than in the first. The differences are small, however, and probably not very significant.

Abelson (1911:1) failed to verify Burt's results but found on

the contrary definite tendencies for correlations to rise with repetition of the tests. Abelson observes with reference to Burt's tests that: " His easier ones may well have been tests of intellectual power to normal children when first tried, but tending to become mechanical in repetition; in that case the more defective children would no longer be at such a disadvantage. Mentally deficient children, on the other hand, would not readily master a performance sufficiently to make it mechanical but would have to continue to exert their full powers " (p. 305).

Simpson (1912:39) compared a group of " good " and a group of " poor " adults with a wide range of tests from motor control to " selective thinking." His main interest was that of a critique of the Spearman theory of a " general factor " and his results have little bearing here. Simpson does comment on the data published by Oehrn and refigured by Spearman and Krüger to the effect " that increase or decrease would depend upon the kind of a test and the stage of the subjects in the learning process . . ." (p. 81).

Colvin (1915:10), Woodrow (1916–7:55), Ordahl and Ordahl (1915:31), and Murdoch (1918:28) have studied the learning of subjects of the same mental ages but of differing chronological ages. Woodrow used the problem of sorting gun wads on which had been pasted various geometrical designs. The following have been selected from his results (Table III, p. 936 and elsewhere):

Group	N.	Aver. M. A.	Aver. Initial Trials	Aver. Final Trials	Aver. Improvement	Aver. per Cent of Improv.
Feebleminded	20	8–10	121	175	55	49
Normal		9–1	122	176	55	46

All of these studies seem to show that the learning curves of subjects of equal mental ages are strikingly similar regardless of the great difference in actual ages, except that of Murdoch, who found that normal children improved more in educational tests over a period of a year than did feebleminded children of the same mental age. But, as L. S. Hollingworth (1920:21, p. 178) has pointed out, Murdoch did not take into account the fact that

"*the children do not remain of equal mental age*" over the period of the experiment, due to the more rapid mental growth of the normal group. Woodrow's experiments do not involve this time error since all of the experimental work was done within a space of about two weeks.

There is almost general agreement that, even where the improvement is equal, the learning curves of the feebleminded show greater daily fluctuations and irregularities than do those of normal children of the same mental ages.

Gould and Perrin (1916:16), using a group of adults and a similar group of children as representing two groups of varying intelligence, compared their curves in maze learning. They concluded that intelligence is manifested chiefly in the initial stages of the learning and that the controlling factors in the later parts of the curves are fatigue and motor control. But they also find that the intelligent learners (the adults) make poorer records on the first two trials than do the less intelligent learners (children) and show a steeper initial rise with greater freedom from steeples. One is tempted to raise the issue here, in view of the results, whether the difference between children and adults in such a function is a matter of intelligence alone and whether such other factors as may be involved hold true explanation of these results. Perrin (1919:32), comparing the learning of adults in the analogies and mirror reading tests, found no correlation between the rankings of the subjects in the two tests, the superior subjects being at their best the farther they were away from the physiological limit of improvement in mirror reading, and the nearer they were to their limit in the analogies test. These relations were reversed in the inferior group. One of Perrin's conclusions is of interest here: "In one respect, the demonstrated lack of relationship is significant. It furnishes justification for the conclusion that the similarity between the tests as regards slope, the greater improvement of the inferior subjects, and the reliability of the initial scores as indices of future accomplishment, *is due to the nature of the tests themselves, rather than to the personnel of*

the practicing group" (pp. 59–60, italics mine). The correlations between initial scores and subsequent improvement were:

Mirror reading..........	.85 to .94
Analogies...............	.59 to .89

Perrin sums up his opinions on the relation of intelligence to learning in the further statement that: ". . . intelligence thus becomes defined in terms of immediate, consistent, and uniform adjustment, not in adjustment considered as a capacity for improvement by leaps and bounds" (p. 51).

Wallin (1916:50) practiced large numbers of children on form boards. The average group improved most, the dullest next, and brightest group somewhat less than the dullest.

Strong (1917:44) says: "The slope of the learning curves of school children based on simple arithmetical combinations apparently correlates to a very considerable extent with the general intelligence of the children" (p. 153). Strong does not, however, present objective data in support of this conclusion.

Myers (1918:30) has already been reported as finding no correlation between the intelligence of normal school children and their abilities in card sorting.

Dallenbach (1919:13) divided a group of feebleminded children into "superior," "medium," and "inferior" groups for study with visual apprehension of numerals, letters, words, and geometrical designs and figures. He used also a group of normal children for purposes of comparison. Mental age correlated with standing in visual apprehension as follows:

Before practice........	0.70 ± .056
After practice.........	0.63 ± .094

He states that: "Individual differences are marked, but they are closely correlated with the mental age" (p. 82).

Johnson (1919:22) studied three groups of five adults each in a target throwing test. The chronological ages of the subjects ranged from about eighteen to twenty-eight years, and the mental ages from about eight to seventeen years. The groups were divided upon the basis of mental age as superior, medium, and

inferior. No correlations are given but it is stated that the superior group had the greatest initial ability as well as the highest final efficiency, the medium and inferior groups following in their respective orders.

Woodrow (1919:56) claims: " What a child can do and how fast he can learn depends upon his mental age " (p. 37–8).

L. S. Hollingworth (1920:21), after reviewing the experimental evidence at some length, supports the contention of Woodrow and others in the following statement: " The feebleminded learn at the same rate, and in the same way as normal children of equal mental age, in tasks in which both have been experimentally tested " (p. 186).

CHAPTER III

Description of the Methods of the Investigation

The Subjects:

In accordance with the experimental aims already recorded, the subjects were selected upon the basis of intelligence or general mental ability. No additional selective agencies are known to have been involved other than those operative on all public school children. Racially the subjects were all of European descent with the exception of one girl who was included in ignorance in advance of the fact that she was partly of negro ancestry.

In all, about 120 different subjects took part in one or more of the experiments. From fifty to sixty-five subjects were used in each experiment. Eleven took part in all three tests. These subjects were pupils in the seventh, eighth, or ninth grades of the University High School, Eugene, Oregon, the Lincoln Elementary School of Oakland, California, or the Oakland High School. The experiments extended over the period from 1919 to 1922.

The chronological ages varied around fourteen years as a mode. The exact ages are recorded in Tables I, II, and III of Chapter IV.[1] Obviously it would have been desirable to have used subjects who were all exactly of the same age but the practical difficulties of securing the requisite numbers at any one age made it necessary to sacrifice to some extent this theoretical advantage and attempt to allow for the influence of the chronological age differences by resorting to the method of partial correlations as an approximation to the results which would have been obtained with a constant chronological age.

The mental age range as shown by the same tables extends

[1] These tables have been omitted from this monograph because of difficulties in printing. They have been bound and placed on file in the Department of Psychology, Stanford University, California, and may be borrowed upon request.

from about eight to eighteen years, thus permitting a range of intelligence quotients of about 75 points. Tables V, VII, and IX of Chapter V present the data on the means and standard deviations of the age groups. The measure of intelligence used, the mental age, was that of the Stanford Revision of the Binet-Simon Scale. In every case at least one standard group test was used as a check on the mental ages obtained by the Binet tests. Where considerable lack of agreement was found, the subject was retested with the Binet tests and the results of the two tests averaged. The mental ages as stated in the tables are always those obtained by the Binet tests and are never based upon a group test alone.

To avoid the problem of the possible influence of sex, the numbers of the sexes were kept roughly equal in all of the experiments.

Description of the Tests:

The three tests which were finally selected in accordance with the criteria already given will be designated as follows:

I. Card Sorting.
II. Code Substitution.
III. Abstract Relations.

A detailed description of each follows:

I. CARD SORTING

The task here consisted in sorting a pack of 100 cards bearing supposedly novel and meaningless designs into a case of ten compartments arranged as two rows of ten compartments each. The following diagram will make this arrangement clear:

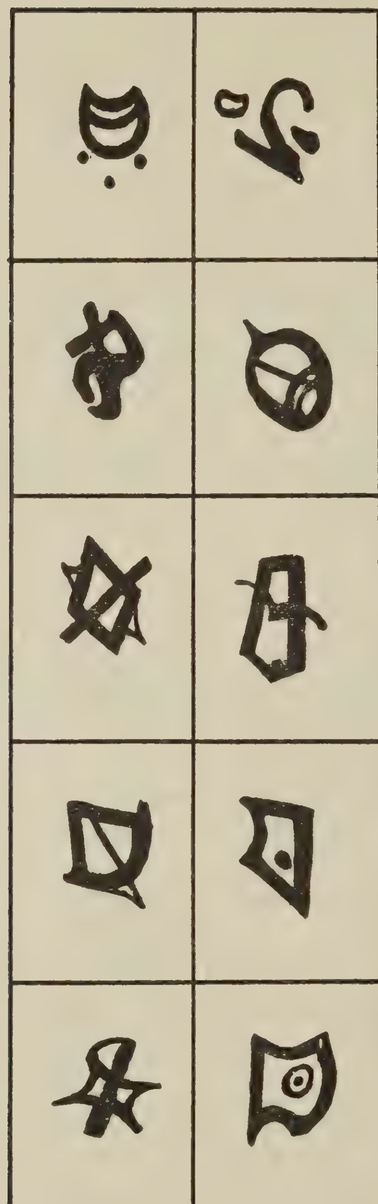

Fig. 1.—Designs on Cards for Sorting Test.

The designs as printed here are full size. The compartments are, of course, greatly reduced. The dimensions of the case were 12 x 28 inches and of the single compartments about 5½ x 5½ inches. The dimensions of the cards were 2½ x 4½ inches.

The pack of cards consisted of ten cards each of ten different suits as determined by the ten designs shown above. The order in the diagram is that of their actual order in the compartments. Each compartment bore a label card corresponding to one of the designs on the cards.

The instructions to the subjects were as given here:

" Here is a pack of 100 cards. Each card has a design printed on it which is probably new to you. There are ten different designs and there are ten cards of each design. (E. shows one of the cards to S. for two seconds.) Under the cover here is a case with ten compartments. Each compartment has a label card and there is a compartment labeled like every kind of design on the cards. (E. raises the cover of the case for two seconds to allow S. a brief inspection.)

" When the signal 'Go!' is given, you are to take the cards one at a time and sort them into the compartments which have the same kinds of label cards. The black line is printed on the cards to show which is the bottom of the card and you should take care to hold them with the black line down. The directions are to go as fast as you can without making mistakes.

" Do you understand? Ready, Go!" (E. removes the cover of the case and at the same time starts the watch. The cover of the case is replaced as soon as the last card is thrown.)

The time was taken to the nearest one-fifth of a second. All errors and changes were noted by the experimenter. The numbers of errors were surprisingly small, averaging fewer than one per trial per subject. Errors were finally ignored for lack of proper method of scoring. The arbitrary practice of adding one second to the time for each error, which has often been used, would not have affected the results to any significant extent.

After the cards were sorted by the subject they were reshuffled by the experimenter in such a way as to avoid the occurrence of two cards of the same suit in consecutive order. Care was also

taken to avoid " runs " of cards in the same order from trial to trial.

Each subject sorted the pack five times each day over a period of ten days (50 trials). The times for the five trials of each day were averaged as the daily score.

An attempt was made to measure the improvement in the mere manipulation of the cards during the ten days of practice by having the subject " box " a similar pack of plain cards five times before the first day's regular practice and five times again at the close of the tenth day's practice. These two sets of boxings were averaged and appear in Table I under the title of " Initial Motor Time " (I.M.T.) and " Final Motor Time " (F.M.T.), respectively. The boxing consisted in throwing the cards one at a time into the compartments taken in the order 1, 2, 3, 4, etc. The difference in the average initial motor time and final motor time was taken as a rough measure of the improvement in the mere manipulation of the cards. The value of this attempted measure will be discussed later.

The practice series was usually continuous except for Sundays. In a few cases both Saturday and Sunday were missed.

Fifty-two subjects took part in this experiment.

II. CODE SUBSTITUTION

The second learning situation was that of transcribing a chapter from Oliver Twist which had been prepared in the code symbols. A second chapter of the same work was provided for translation into the code symbols. Each subject was given a key card bearing the code. This was kept before the subject during the practice until such a time as its use was voluntarily abandoned. The key to the code is given herewith.

The daily practice consisted of the translation for ten minutes from the code into English, followed by translation for the same amount of time from English into the code. Two minutes rest period was allowed between the two exercises. The total practice time was therefore twenty minutes a day.

The score for each sort of translation was the number of letters

(or symbols) transcribed correctly in the ten minutes. No additional penalty was added for mistakes other than loss of credit for that letter or symbol.

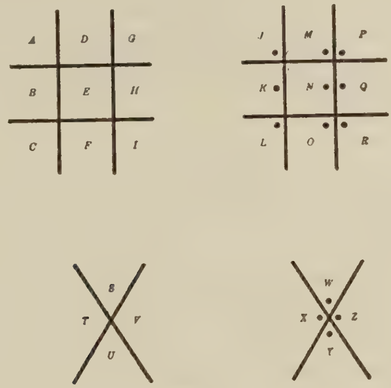

Fig. 2.—Key for Code Substitution Test.

Sixty-six subjects took part in this experiment. Twenty-one continued after the regular ten day period for varying numbers of days.

Samples of the test material will be found in the Appendix.

The exact instructions to the subjects were as follows:

" This test is one in which you have to translate a story written in a code or secret language into English words. You may have seen the code before in some of the tests which we have given. Whether you have seen the code before does not matter, as you will have a copy of the code before you as you work.

" Now look at the key cards which I have given you and we will write a few words together so that you will understand how to use the code when we turn over the page and begin the real work. You watch the code as I write some words on the blackboard. (E. writes ' University of Oregon ' slowly, pointing out each symbol on the key card.)

" When the signal to turn is given, turn over the page and begin with the first code sign and write the English letter directly under it. Take the others in order and be sure that you know what each word is before you go on to the next. In this way the meaning of the story will often help you in the translation of the code.

"I am going to give you exactly ten minutes to work and I want to see how many words you can do in that time without making mistakes. **Both speed and accuracy are important.** The real purpose of this test is to see how much improvement you can make each day. For these reasons you should work just as hard as you can to make a good record every day just as if you were playing a game. **Remember, both speed and accuracy are important.**

"Ready, turn, Go!"

The following additional directions for the English code translation were needed:

"Here you have to write the code word under the printed word. Since you cannot easily make the code letters as small as the printed letters, the lines only go half way across the page. You can write straight across the page beginning the line at the left. Ready, Go!"

The code test was used as a semi-group test. The instructions and first day's practice were given individually. Sometimes the subjects worked alone for two or three days. Usually, however, the subjects were formed into small groups of a half dozen after the first day. Judging from careful inspection this plan sacrificed little or nothing in the validity of the results, since any disturbing factors brought about by working in small groups were wholly or nearly wholly compensated for by the added stimulation of group competition. The experimenter checked up on this point by questioning the more intelligent subjects. In fact, it became the usual practice for the subjects to get together after the practice period and compare notes on their respective gains.

III. ABSTRACT RELATIONS TEST

The administration and nature of this test will be made clear by the instructions to the subjects which are given at length here. The introductory directions and explanations were necessarily very detailed and lengthy, due to the fact that the use of subjects with mental ages of eight or nine years made it imperative that every precaution be taken to insure the initial comprehension of

the task. The hope was that failure to make a positive performance in the test would indicate inability to solve the test problems after the instructions were understood and not merely lack of comprehension of these instructions.

A key card was given to each subject and this was kept before him until voluntarily discarded. The key gives the clue to the nature of the problem.

Key

" C IS LARGER THAN A
A IS LARGER THAN B
B IS LARGER THAN D "

" If we arrange the letters in the order of the largest to the smallest, they are: C–A–B–D.

Samples

1	— C plus D equals B	(E. explains)
2	+ A minus B is less than C minus B	"
3	± C minus B is larger than D	"
4	— A minus D is equal to C	"
5	± B plus C plus D is larger than C plus A	"
6	— A plus D is less than B minus C	"
7	— B minus A is less than D minus C	"

" The above samples are already marked correctly. Now begin with problem one of the regular test sheet and work the problems in order. *Be sure that you have answered each problem correctly before you go on to the next one.* Mark those that are *always untrue* with a minus sign (—), those that are always true with a plus sign (+), and those that might or might not be true with both a plus and a minus sign (±). You will be allowed to keep the key card before you at all times as you work. You will be told to stop at the end of fifteen minutes. Ready, Go! "

In order to make as clear as possible the task to be performed, the experimenter demonstrated the solution of the seven samples on the key card as just given. E. said (verbatim):

" You have noticed from the key card that C is larger than A; A is larger than B; B is larger than D, so that if we arrange the letters in the order of the largest to the smallest, they are: C. . . . A. . . . B. . . . D. . But, you must always remember that

you do not know how much larger C is than A, or A is than B, or B is than D. All that you know is that C is larger than A, and A is larger than B, and B is larger than D.

"All of the letters stand for numbers that are more than zero. However, some of these problems do involve combinations of numbers that are equal to or even less than zero. Examples of these will be explained to you later.

"There is one other point which you must remember or you will make mistakes. It is very important that you do not try to work these problems by letting numbers stand for the letters. These four letters do not stand for any certain numbers. For example, if A were equal to 10, then C might be 11 in one problem and more than a million in the next. The same thing is true of all the letters. You do not know how large they are but only that they stand in certain order to each other as the key card states. **You are sure to make mistakes if you try to substitute actual numbers for the letters.**

"Now look at sample 1. C plus D equals B. We would reason out the answer to this problem like this: since C is the largest of all the letters, it must be larger than B. If, then, we add D to C, we would make C still larger than B. The statement that C plus D equals B is always untrue, and it has been marked with a minus sign in front of it. The minus sign means that the statement never could be true. Do you understand? (The explanation is repeated in case the subject fails to comprehend.)

"*Sample 2.* A minus B is less than C minus B. In this sample problem you will notice that the same number is to be subtracted from both sides of the problem. What number is it? (Pause to see whether S. answers. In case S. does not, E. continues.) B is to be subtracted from both sides. Since this is true, we can ignore the B and we have left the statement that A is less than C. This is always true and it has been marked with a plus sign.

"*Sample 3.* C minus B is larger than D. We know that C is always larger than D, but, if B were large enough, C minus B might be less than D. If, however, B is very small, C minus B might still be larger than D. Since we don't know how large any of the numbers are, we can only say that the statement may or may not be true. It is marked therefore with both a plus and a minus sign.

"*Sample 4.* A minus D is equal to C. We know that A is less than C. If now we subtract D from A it will be still smaller

than C. The statement must always be untrue and is marked minus.

"*Sample 5.* B plus C plus D is larger than C plus A. In this problem which letter is found on both sides? (Pause as before for answers.) Since we are adding C to both sides we can cancel out the C and have left B plus D is larger than A. Both B and D are smaller than A. But, the sum of the two may be either larger, equal to, or smaller than A. We have marked it plus and minus because it may or may not be true.

"*Sample 6.* A plus D is less than B minus C. Notice the expression B minus C on the right-hand side. Since C is larger than B, this must equal less than zero. A plus D is more than zero and hence must be larger than B minus C. It is therefore always untrue and has been marked minus.

"*Sample 7.* B minus A is less than D minus C. This problem is very much like the one before except that both sides are equal to less than zero. It is very much harder to decide which is the larger when both are less than zero. Numbers less than zero are called minus numbers and are something like debts. If Mr. A. has $150 and Mr. B. has $20, which is the richer man? (S. answers and E. accepts or corrects as before.) Now, if instead of having these sums of money, A. owes $150 and B. owes $20, which is the richer man? (S. answers as before.) This is why minus numbers are like debts, the larger a minus number appears to be, the smaller it really is. Minus 100 is less than minus 10, and minus 10 is less than minus 1, and minus 1 is less than zero. Look again at sample 7. B minus A is less than D minus C. D is the smallest number of all and C is the largest. If we take the largest of all from the smallest of all, we have left a bigger minus number than if we took the next to the largest (A) away from the next to the smallest (B). But like a debt, the bigger a minus number looks to be, the smaller it really is. B minus A is larger, then, than D minus C and the statement is untrue. (The explanation is repeated once more if asked for. However, no further efforts to explain are made.)

"Now look at the directions at the bottom of the key sheet. You are to reason out these problems one at a time and mark those that are always true with a plus sign, those that are always untrue with a minus sign, and those that might or might not be true with both signs.

"Remember two things: (1) Be sure that you are right before you go on to the next problem, and (2) do not substitute

numbers in place of the letters as you are sure to make mistakes. Ready, turn, go!"

At first glance the instructions appear to be exceedingly involved. However, this was absolutely necessary in order to give the dullest subjects every possible opportunity to make progress in the test. Zero scores could then be interpreted as failures to achieve rather than failure to comprehend instructions. The brighter subjects understood the instructions readily, or were even somewhat amused at the efforts to explain what to many of them was very obvious. A few of the subjects had studied algebra. The inclusion of a few subjects with mental ages below ten years who made zero scores was intentional in order that a close approach to the zero point in ability in this test might be had. In this case it should be pointed out that six subjects who failed to make better than a zero average score for the first five days were allowed to discontinue on the sixth day and their scores for all ten days recorded as zero.

The daily time limit was fifteen minutes. The test material for a given day consisted of 100 problems of the types represented in the seven samples (see also in the Appendix). Three approximately equivalent sets of these were provided. Set I was given on the first, fourth, seventh, and tenth days. Set II on the third, sixth, and ninth days. In case a subject continued longer than ten days, the same order of rotation of the forms was followed.

Because the number of possible responses was limited to three, a correction for chance was necessary. This was done by the method of scoring the number right minus one-half the number wrong. Omissions were not counted either way. In case a subject completed the entire 100 problems of a set in less than the 15 minute time limit, his actual score was multiplied by $\frac{15}{\text{actual working time}}$. This is termed his "rate score." His "accuracy score" was obtained by dividing the number right corrected by the number attempted. An example will make the scoring

clear. Suppose a subject finished a series of 100 problems in 12 minutes and 15 seconds with six omissions and ten mistakes:

Series.	100
Omissions.	6
Attempts.	94
Wrong.	10
Right.	84
R—W/2	79

$$\text{Rate score} = 79 + \frac{15}{12.25} = 96.7$$

$$\text{Accuracy score} = \frac{79}{94} = 84.0$$

The speed score involves the assumption that the subject would have completed the estimated number had he continued for 15 minutes on similar materials at the same rate of accuracy.

Measured in time units, those subjects who finished in less than the time limit were less practiced at the end of the experiment than the slower subjects. In terms of the number of problems attempted, the rapid workers were more practiced at the end than the slower ones. Since so little is known of the comparative merits of time versus work units, it is wholly conjectural which group of subjects was favored by the plan adopted.

More than sixty subjects took part in this experiment for at least ten days. The method was that of the semi-group plan which was described in the discussion of the code-substitution test.

The subjects were not told their daily scores in any of the tests. If inquiries were made, answers of general encouragement were given. However, many of the subjects spontaneously tried to check up their gains by noting the number of test items attempted from day to day. No efforts were made to prevent the subjects comparing notes on their gains in this way.

CHAPTER IV

The Experimental Results

The results of the three experiments have been recorded in Tables I, II, and III. Because of difficulties in printing, these tables have been suppressed in this report and have been placed on file in the Department of Psychology of Stanford University, and may be borrowed upon written request. A brief characterization of each table is given below.

Table I gives the results for the fifty-two subjects used in card sorting. The daily scores are given as the averages of the five daily trials. Averages for the ten days of practice are also given, such averages including the entire fifty trials. Averages of five preliminary and five final " boxings "' are also recorded. These have been described elsewhere. Table I–A, a supplement to Table I, shows the time scores for the first five trials separately, *i.e.,* the five trials entering in to the average score of the first day.

Table II presents the results for the code-substitution experiment. Daily scores are given in terms of the numbers of letters (or symbols) translated in ten minutes. Two figures are given for each day, the upper value being the code to English score and the lower the English to code score. A supplement to Table II, designated as Table II–A, gives the scores for a number of subjects who continued practice more than 15 days, in some cases as long as 40 days.

Table III presents the scores in the abstract relations test. Two daily scores are given for each subject, the upper being the " accuracy " score and the lower the " rate " score. The description of the computation and meaning of these measures has already been given in Chapter III.

It is to be regretted that the full tabular statement of the original scores cannot be reproduced here. The reader will, however, experience little difficulty in following the statistical treatment, since the tables of Chapter V present the means and standard deviations of all subjects for all of the learning tests as well as for chronological age, mental age, etc.

CHAPTER V

Statistical Treatment and Interpretation of the Results

Two general methods were available in the treatment of the experimental data in order to reveal the part played by general intelligence in the rate of learning. The first of these is that of a division of the practicing group into two or three sub-groups upon the basis of the mental ages, *e.g.*, into a " superior," an " average," and an " inferior " group. The same sort of a division might also be made upon the basis of the intelligence quotients, *e.g.*, those above 110, those from 90 to 100, and those below 90. The former has the advantage that it would be less open to the objection that the chronological ages of the subjects are not constant and hence the I.Q.'s alone would introduce a variable factor from group to group.

The second method which was open to use in the present study allows of even greater refinement in the corrections for the varying chronological ages of the subjects. If the daily performances are correlated against mental ages, it is possible to eliminate the influence of chronological age by the use of the method of partial correlations. This necessitates the computation of all the possible intercorrelations between the three variables, viz., mental age, performance, and chronological age. Substantially the same facts will be revealed by each method. The first has the advantage of being adapted to presentation in graphic form, the second of greater statistical refinement. For these reasons both methods will be used although the main treatment of the results will be based upon the method of partial correlations.

In the computation of the coefficients of correlation the Pearson product-moment formula was used in all cases. One general form of this formula is:

$$r = \frac{\Sigma x y}{N \sigma_1 \sigma_2}$$

The probable errors of these correlations were obtained by the formula:

$$P.E._r = .6745 \frac{1-r^2}{\sqrt{N}}$$

The formula for the partial coefficients of correlation follows:

$$r_{12\cdot3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1-r^2_{13}}\,\sqrt{1-r^2_{23}}}$$

In order to secure a check upon the changes in the relative variabilities of the performances of the practicing group from time to time, the Pearson Coefficients of Variation have been computed for each day's performance. These coefficients state the relative variabilities in terms of the ratio of standard deviations to the means. The standard deviation alone is a measure of the variability in terms of gross scores and hence is dependent upon the numerical magnitudes of the units of measurement. The formula for the Pearson Coefficient of Variation is:

$$V = \frac{100\ \sigma}{\text{Mean}}$$

Since the relative heterogeneity of the talent as represented by the daily scores is subject to change, such coefficients are of value in interpreting changes in the magnitudes of the correlation coefficients.

As was stated in the introductory chapter, the selection of the tests finally used in the present study was an attempt to sample a wide range of possible learning situations. For this reason the initial correlations between performance and mental age present some interest as a check upon the wisdom of the original selections. Table IV gives such correlations which have been selected from the larger tables which follow (Tables V, VII, and IX).

TABLE IV

Closeness of Relationship Between the Tests Used and General Intelligence.

Test	$r_{12 \cdot 3}$	P.E.
I Card sorting:		
Trial 1, first day	0.329	.083
Average of 5 trials, first day	0.176	.091
II Code substitution:		
Code to English, first day	0.699	.043
English to code, first day	0.666	.046
III Abstract relations:		
Rate, first day	0.730	.040
Accuracy, first day	0.800	.031

The notation "$r_{12 \cdot 3}$" refers here, as always, to the partial correlation between mental age and performance when the influence of chronological age is eliminated.

It will be seen from this table that ability to sort cards involves relatively little of the ability measured by tests of general intelligence. The mental processes involved in substitution abilities are much more closely related to general mental ability, as is shown by the moderately high correlations obtained. The mental functions involved in the solution of the problems of the abstract relations experiment can be said to be rather closely related to general intelligence as shown by the correlation of $0.800 \pm .031$. These tests, then, do in some measure sample the range of mental functions with respect to the demands made upon intelligence from functions of very low relationship to functions of fairly close identity. It must be remembered that in view of the very large range of talent employed (see Tables VI, VIII, and X for the standard deviations of the mental ages), the exact magnitudes of these correlations have little significance.

The chief interest for present purposes is not concerned with the question of the exact magnitudes of correlations of performance and mental age, but rather with the evidence of *systematic tendencies toward either increase or decrease of such correlations with practice.* Tables V, VII, and IX present the partial coefficients and the intercorrelations between the three variables for

each test. The notations of the variables by number are invariably as follows:

1 Mental age.
2 Performance.
3 Chronological age.

Tables VI, VIII, and X give the means, standard deviations, and Pearson Coefficients of Variation for the separate variables in the different tests.

Table XI gives the correlations between initial and final performance in all of the tests. Initial performance refers to the score on the first day of practice and final performance refers to score on the tenth day of practice.

In order to interpret certain details of the experimental results which apply only to individual tests, each of the experiments will be discussed separately.

TABLE V

Correlations for Card Sorting

1st 5 Trials i.e., first day of Practice	r_{12} M.A. and Score	r_{13} M.A. and C.A.	r_{23} C.A. and Score	$r_{12\cdot3}$ M.A. and Score Independent of C.A. (partial)	N
1	0.434 ± .076	—0.379 ± .080	—0.412 ± .078	0.329 ± .083	52
2	0.403 ± .078	—0.379 ± .080	—0.454 ± .074	0.280 ± .086	52
3	0.192 ± .090	—0.379 ± .080	—0.348 ± .082	0.069 ± .093	52
4	0.219 ± .089	—0.379 ± .080	—0.278 ± .086	0.128 ± .092	52
5	0.211 ± .089	—0.379 ± .080	—0.309 ± .085	0.107 ± .092	52
Days of Practice, i.e., averages 5 daily trials					
1	0.310 ± .085	—0.379 ± .080	—0.430 ± .076	0.176 ± .091	52
2	0.058 ± .093	—0.379 ± .080	—0.402 ± .078	—0.112 ± .092	52
3	0.015 ± .093	—0.379 ± .080	—0.305 ± .085	—0.114 ± .092	52
4	0.038 ± .093	—0.379 ± .080	—0.314 ± .084	—0.092 ± .093	52
5	0.052 ± .093	—0.379 ± .080	—0.175 ± .091	—0.015 ± .094	52
6	0.014 ± .096	—0.388 ± .082	—0.188 ± .093	—0.066 ± .096	49
7	—0.016 ± .096	—0.388 ± .082	—0.153 ± .094	—0.083 ± .096	49
8	—0.077 ± .097	—0.365 ± .084	—0.201 ± .093	—0.164 ± .095	48
9	—0.035 ± .096	—0.365 ± .083	—0.195 ± .093	—0.117 ± .095	49
10	—0.060 ± .096	—0.367 ± .083	—0.041 ± .096	—0.081 ± .096	49

Variables: 1. Mental Age.
2. Score in Seconds.
3. Chronological Age.

TABLE VI

Mean Scores, Standard Deviations, and Coefficients of Variation in Card Sorting

Trials 1–5, i.e., first day of Practice	Mean Mental Age	S.D. Mental Age in Months	Mean Chron. Age	S.D. Chron. Age in Months	Mean Score in Sec.	S.D. Score in Sec.	Pearson Coeff. of Variation	N
1	13—9.1	26.9	13—7.7	13.7	293.5	76.1	25.93	52
2	13—9.1	26.9	13—7.7	13.7	195.5	49.7	25.42	52
3	13—9.1	26.9	13—7.7	13.7	162.5	35.8	22.03	52
4	13—9.1	26.9	13—7.7	13.7	149.0	31.9	21.41	52
5	13—9.1	26.9	13—7.7	13.7	140.8	26.7	18.96	52
Days of practice								
1	13—9.1	26.9	13—7.7	13.7	186.3	37.6	20.19	52
2	13—9.1	26.9	13—7.7	13.7	111.3	24.7	22.22	52
3	13—9.1	26.9	13—7.7	13.7	104.7	21.2	20.27	52
4	13—9.1	26.9	13—7.7	13.7	94.7	23.1	24.39	52
5	13—9.1	26.9	13—7.7	13.7	97.9	21.0	21.42	52
6	13—9.1	27.2	13—8.0	14.1	96.2	22.0	22.86	49
7	13—9.1	27.2	13—8.0	14.1	93.6	20.3	21.74	49
8	13—9.9	27.0	13—7.6	14.0	91.7	19.7	21.47	48
9	13—10.2	26.9	13—7.5	13.8	92.4	22.2	24.00	49
10	13—10.2	26.9	13—7.5	13.8	93.3	19.9	20.31	49

TABLE VII

Correlations for Code Substitution

Day of Practice	r_{12} M.A. and Score	r_{13} M.A. and C.A.	r_{23} C.A. and Score	$r_{12\cdot3}$ M.A. and Score Independent of C.A.	N
Code to English					
1	0.698 ± .043	—0.044 ± .083	—0.129 ± .082	0.699 ± .043	66
2	0.633 ± .050	—0.044 ± .083	—0.028 ± .083	0.633 ± .050	66
3	0.670 ± .046	—0.044 ± .083	—0.010 ± .083	0.670 ± .046	66
4	0.635 ± .049	—0.044 ± .083	—0.019 ± .083	0.635 ± .049	66
5	0.613 ± .052	—0.044 ± .083	—0.012 ± .083	0.613 ± .052	66
6	0.631 ± .050	—0.044 ± .083	—0.079 ± .082	0.630 ± .050	66
7	0.607 ± .052	—0.044 ± .083	—0.084 ± .083	0.613 ± .052	66
8	0.683 ± .044	—0.044 ± .083	0.041 ± .083	0.686 ± .044	66
9	0.613 ± .054	—0.023 ± .086	0.047 ± .086	0.615 ± .054	61
10	0.608 ± .056	—0.049 ± .088	0.054 ± .088	0.613 ± .055	58
English to Code					
1	0.666 ± .046	—0.044 ± .083	—0.020 ± .083	0.666 ± .046	66
2	0.629 ± .050	—0.044 ± .083	—0.081 ± .082	0.628 ± .050	66
3	0.598 ± .053	—0.044 ± .083	—0.057 ± .083	0.597 ± .053	66
4	0.520 ± .060	—0.044 ± .083	—0.028 ± .083	0.520 ± .060	66
5	0.513 ± .061	—0.044 ± .083	—0.075 ± .083	0.511 ± .061	66
6	0.531 ± .060	—0.044 ± .083	—0.036 ± .083	0.530 ± .060	66
7	0.607 ± .052	—0.044 ± .083	—0.010 ± .083	0.607 ± .052	66
8	0.582 ± .055	—0.044 ± .083	0.011 ± .083	0.583 ± .055	66
9	0.519 ± .063	—0.023 ± .086	0.027 ± .086	0.520 ± .060	61
10	0.576 ± .059	—0.049 ± .088	—0.029 ± .088	0.576 ± .066	58

TABLE VIII

Mean Scores, Standard Deviations, and Coefficients of Variation in Code Substitution

Day of Practice	Mean Mental Age	S.D. Mental Age in Months	Mean Chron. Age	S.D. Chron. Age in Months	Mean Score (Lett.)	S.D. Score (Lett.)	Pearson Coeff. of Variation	N
Code to English								
1	14—3.6	34.9	14—1.1	15.0	94.09	50.48	53.65	66
2	14—3.6	34.9	14—1.1	15.0	157.88	62.85	39.81	66
3	14—3.6	34.9	14—1.1	15.0	197.27	69.88	35.42	66
4	14—3.6	34.9	14—1.1	15.0	208.48	65.63	31.48	66
5	14—3.6	34.9	14—1.1	15.0	222.12	66.53	29.95	66
6	14—3.6	34.9	14—1.1	15.0	236.36	71.77	30.36	66
7	14—3.6	34.9	14—1.1	15.0	247.27	77.45	31.32	66
8	14—3.6	34.9	14—1.1	15.0	255.45	90.22	35.32	66
9	14—4.8	35.0	14—1.3	15.1	270.33	89.04	32.94	61
10	14—4.3	35.8	14—1.0	15.1	278.28	90.59	32.55	58
English to Code								
1	14—3.6	34.9	14—1.1	15.0	110.68	55.88	50.49	66
2	14—3.6	34.9	14—1.1	15.0	153.64	54.71	35.61	66
3	14—3.6	34.9	14—1.1	15.0	176.67	61.92	35.05	66
4	14—3.6	34.9	14—1.1	15.0	193.33	62.53	32.34	66
5	14—3.6	34.9	14—1.1	15.0	197.58	61.37	31.06	66
6	14—3.6	34.9	14—1.1	15.0	208.79	67.77	32.46	66
7	14—3.6	34.9	14—1.1	15.0	218.48	75.34	34.48	66
8	14—3.6	34.9	14—1.1	15.0	227.58	83.32	36.61	66
9	14—4.8	35.0	14—1.3	15.1	242.79	89.59	36.90	61
10	14—4.3	35.8	14—1.0	15.1	255.17	91.00	35.66	58

I. CARD SORTING

Table V shows definitely that intelligence[1] offers but slight and temporary advantage in learning to sort cards. If the first five trials which make up the first day's practice are alone considered, it appears that, even in the course of these five trials, the partial coefficients (column headed $r_{12 \cdot 3}$) decreased from 0.329 ± .083 for the first trial to 0.107 ± .092 for the fifth trial. The value for the third trial is even lower, being 0.069 ± .093. If the daily averages of the five trials are compared, it will be seen that these values fell from 0.176 ± .091 on the first day to —0.112 ± .092

[1] In the interests of concise expression, the term "intelligence" has been used throughout this discussion as if it were a unitary or elemental factor or mental function. In reality no such definition is implied. For present purposes, no other definition is offered or needed than to point out that intelligence is here defined in terms of what the Binet tests really measure. Intelligence, therefore, is synonymous with Binet mental ages.

on the second day. The coefficients on subsequent days remained slightly negative but not significantly so in view of the large probable errors.

The changes in the relation of chronological age to performance is even more marked in degree. Starting with a correlation of —0.430 ± .076 on the first day, the coefficients approach zero more and more closely each day, the final value being —0.041 ± .096. In view of the negative correlation between the M.A. and the C.A. of the entire group (—0.379 ± .080), the two sets of coefficients are quite in harmony with each other.

One general fact seems evident throughout all of the correlations presented for card sorting, viz., that at the end of practice the subjects have arranged themselves within the group in an order which is quite independent of either their mental maturity or their age. This applies, of course, only within the limits of the age range represented in the experiment. Performance in card sorting, it can be concluded, is controlled by factors of a

TABLE IX

Correlations for the Abstract Relations Test

Day of Practice	r_{12} M.A. and Score	r_{13} M.A. and C.A.	r_{23} C.A. and Score	$r_{12\cdot3}$ M.A. and Score Independent of C.A.	N
Rate					
1	0.720 ± .041	0.092 ± .084	—0.060 ± .085	0.730 ± .040	63
2	0.804 ± .030	0.092 ± .084	0.079 ± .084	0.803 ± .030	63
3	0.819 ± .028	0.092 ± .084	0.053 ± .085	0.818 ± .028	63
4	0.786 ± .032	0.092 ± .084	0.034 ± .085	0.787 ± .032	63
5	0.814 ± .029	0.092 ± .084	0.066 ± .085	0.813 ± .029	63
6	0.789 ± .032	0.092 ± .084	0.096 ± .084	0.787 ± .032	63
7	0.802 ± .030	0.092 ± .084	0.087 ± .084	0.801 ± .030	63
8	0.815 ± .028	0.092 ± .084	0.107 ± .084	0.813 ± .029	63
9	0.813 ± .029	0.095 ± .085	0.076 ± .085	0.811 ± .029	62
10	0.820 ± .028	0.095 ± .085	0.095 ± .085	0.819 ± .028	62
Accuracy					
1	0.793 ± .032	0.092 ± .084	—0.034 ± .085	0.800 ± .031	63
2	0.712 ± .042	0.092 ± .084	0.008 ± .085	0.714 ± .042	63
3	0.748 ± .037	0.092 ± .084	0.035 ± .085	0.748 ± .037	63
4	0.817 ± .028	0.092 ± .084	—0.048 ± .085	0.826 ± .027	63
5	0.829 ± .027	0.092 ± .084	0.017 ± .085	0.831 ± .026	63
6	0.833 ± .026	0.092 ± .084	0.015 ± .085	0.835 ± .026	63
7	0.766 ± .035	0.092 ± .084	0.035 ± .085	0.766 ± .026	63
8	0.837 ± .025	0.092 ± .084	0.043 ± .085	0.837 ± .025	63
9	0.822 ± .028	0.095 ± .085	0.030 ± .086	0.823 ± .028	62
10	0.842 ± .025	0.095 ± .085	0.006 ± .086	0.845 ± .025	62

TABLE X

Mean Scores, Standard Deviations, and Coefficients of Variation for Abstract Relations

Day of Practice	Mean Mental Age	S.D. Mental Age in Months	Mean Chron. Age	S.D. C.A. in Months	Mean Score in Problems	S.D. Score in Problems	Pearson Coeff. of Variation	N
Rate								
1	13—9.1	33.5	13—6.6	7.05	20.04	15.48	77.25	63
2	13—9.1	33.5	13—6.6	7.05	35.75	25.99	72.70	63
3	13—9.1	33.5	13—6.6	7.05	40.91	27.51	67.25	63
4	13—9.1	33.5	13—6.6	7.05	42.54	31.84	74.85	63
5	13—9.1	33.5	13—6.6	7.05	49.56	37.22	75.10	63
6	13—9.1	33.5	13—6.6	7.05	55.48	37.52	67.63	63
7	13—9.1	33.5	13—6.6	7.05	58.81	43.48	73.93	63
8	13—9.1	33.5	13—6.6	7.05	65.95	44.50	67.48	63
9	13—9.0	33.7	13—6.7	7.06	68.39	47.48	69.43	62
10	13—9.0	33.7	13—6.7	7.06	69.68	49.28	70.72	62
Accuracy								
1	13—9.1	33.5	13—6.6	7.05	40.12	24.65	61.44	63
2	13—9.1	33.5	13—6.6	7.05	50.83	29.65	58.33	63
3	13—9.1	33.5	13—6.6	7.05	52.98	29.09	54.91	63
4	13—9.1	33.5	13—6.6	7.05	48.33	29.15	60.31	63
5	13—9.1	33.5	13—6.6	7.05	51.67	31.06	60.11	63
6	13—9.1	33.5	13—6.6	7.05	51.43	29.24	56.85	63
7	13—9.1	33.5	13—6.6	7.05	49.40	31.13	63.02	63
8	13—9.1	33.5	13—6.6	7.05	51.19	29.72	58.06	63
9	13—9.0	33.7	13—6.7	7.06	51.65	30.39	58.84	62
10	13—9.0	33.7	13—6.7	7.06	50.02	31.64	63.25	62

specific nature not related to general intelligence. That these specific factors show some tendency to be stable throughout the course of the learning is shown by the correlation of 0.516 ±.071 (Table XI) between initial and final performance. This finding is quite in accord with the results of Wells, Hollingworth, Whitely, Chapman, Terman, Perrin, *et al.,* previously cited.

The transitory initial correlation of efficiency and intelligence presents especial interest because of the attempts of Gould and Perrin *et al.,* to define intelligence in terms of rapid initial adjustment to a learning situation. In Chapter II these authors were quoted as finding that in maze learning differences in the performances of groups of varying intelligence were chiefly in evidence in the initial stages of the learning. These investigators found that their intelligent group (adults) made poorer initial records than the less intelligent group (children). Such findings are exactly opposed to the results given here for card sorting. It

should also be objected that their use of adults and children as groups differing in intelligence alone, without reference to the other factors which must be involved in maze learning, is probably an unwarranted procedure.

During the practice period, as shown by Table VI, the standard deviations and the means of the time scores are reduced to about one-half of the initial size. This, however, does not involve any marked changes in the relative variabilities of the performances as is shown by the column of Pearson coefficients of variation, the first and tenth days giving almost identical values

TABLE XI

Correlations Between Initial Performance (i.e., first day) and Final Performance (i.e., tenth day) for Each of the Three Tests

	Test	r	P.E.
I	Card Sorting	0.516	.071
II	Code Substitution:		
	Code to English	0.606	.056
	English to Code	0.779	.035
III	Abstract Relations:		
	Rate	0.782	.033
	Accuracy	0.804	.030

at 20.19 and 20.31, respectively. Change in the relative variability of the practicing group, therefore, cannot explain the changes in the magnitudes of the correlations which have been noted.

In order to determine, if possible, whether the absence of correlation between final performance and mental age was due to the fact that the subjects were nearing their limits of improvement and consequently the time scores were becoming more and more measures of the time given over to the purely motor processes of throwing the cards into the compartments, one further set of computations was attempted. The average time required for "boxing" the cards five times on the tenth day of practice was subtracted from the average regular sorting time of that day. This difference, it was hoped, might be a rough measure of any elements in the total mental processes involved other than the purely motor functions. It was thought that any remaining sensory processes in the learning which had been obscured by the

motor factors might thus be separated out from the total time. The difference between the final sorting time and the "boxing" time has been termed "Difference Time." Correlations of this time with mental age and the partial correlations of these two variables excluding chronological age are given in Table XII.

TABLE XII

Correlations of Intelligence and "Difference Time"

	1 M.A.	2 "D.T."	3 C.A.
1 M.A.	....	0.055 ± .096	—0.367 ± .083
2 "D.T."	0.055 ± .096	....	—0.149 ± .094
3 C.A.	—0.367 ± .083	—0.149 ± .094	....

$r_{12 \cdot 3} = 0.0004 \pm .0964$

It will be seen at once that this "Difference Time" proved of no significance as far as its relation to intelligence is concerned. This conclusion is identical with that of Brown (1914:6), who reported a similarly attempted measure as invalid.

II. CODE SUBSTITUTION

The situation with respect to the code test is quite different from that in card sorting. Tables VII and VIII give the data on correlations, central tendencies, and variabilities. In both types of substitution, *i.e.*, code to English and English to code, we find moderately high initial correlation with mental ability. The partial coefficients on the first day are 0.699 ± .043 and 0.666 ± .046, respectively. Both sets of correlations drop slightly during the course of practice. Nevertheless, the decrease is probably not very significant in comparison with the probable errors.[2] The partial correlations on the tenth day are 0.613 ± .055 for code into English and 0.576 ± .060 for English into code translation, the correlations are even lower on intermediate days, *e.g.*, 0.511 ± .061 on the fifth day and 0.520 ± .060 on the fourth day.

[2] By means of the formula for the probable error of a difference, it can be shown that obtained differences in order to be very significant would have to be as large as .25 or larger. This formula reads:

$$\text{P.E. (Diff.)} = \sqrt{\text{P.E.}^2_1 + \text{P.E.}^2_2}.$$

This will give values for the P.E. of the difference around .085 where the P.E.'s of the r's are .06. The usual interpretation is that of calling differences uncertain unless the difference is at least 3 times its probable error.

The decrease in the magnitude of the correlations between performance and intelligence is relatively slight and the indications are that the onset of automatization of the mental processes involved would be very much delayed in comparison with card sorting.

The question might be raised again here whether the fall in the size of the correlations (which seems to be most marked from the first to the second day) is merely the result of the decreased variability of the performances of the subjects on the second day as compared with the first, as shown in Table VIII. It is true that the relative variabilities do decrease sharply from the first to the second day, viz., from 53.65 to 39.81 for code to English and from 50.49 to 35.61 for English to code. It is further to be noted that after the second day there is comparatively little change in the relative variability of the subjects. Against this possibility can be urged the fact that the correlations are practically as high on the third and eighth days for code to English as they were at the outset, and yet their Coefficients of Variation are smaller in about the ratio of 53:35. Similar results will be found to hold true of English to code translation, the fluctuations in the sizes of the Pearson Coefficients of Variation seem to be little related to fluctuations in the correlation coefficients. Most of these minor fluctuations are not significantly large in comparison with the probable errors of the coefficients of correlation.

Attention should be drawn to the fact that the range of talent is greater in the subjects practicing with the substitution test than was the case in the card-sorting test, the standard deviations of the mental ages being, respectively, about 35 and 27 months. This difference in the spread of the talent had operated naturally to increase the relative difference between the amount of correlation found between intelligence and card sorting, on the one hand, and intelligence and substitution ability on the other. The range of talent in the subjects used for card sorting is probably not very dissimilar to that ordinarily found within a single school grade but is much greater than that of a single grade for the substitution test.

As shown by the correlations between initial and final ability in the code test (Table XI), the subjects maintain their relative positions with great fidelity, the correlations being $0.606 \pm .056$ for code to English and $0.779 \pm .035$ for English to code. This constancy is considerably greater than was found for card sorting.

In general the fall in the correlations during practice is greater in the case of the English to code than in the reverse type of translation. The reason for this is not entirely evident. It might be attributable to the influence of the value of the use of the context in translation of code into English, upon the assumption that the context is a function analogous to comprehension ability in reading and hence offers greater opportunity for the operation of intelligence, or it may be purely a chance phenomenon. There is no experimental evidence which can be cited in proof or disproof of this hypothesis.

III. ABSTRACT RELATIONS TEST

In this test we find a different situation with respect to the influence of intelligence than in either of the two preceding experiments. As shown by Table XI, the initial correlations between mental age and performance are high; the partials excluding the influence of chronological age differences being $0.730 \pm .033$ for rate and $0.800 \pm .030$ for accuracy.

In the main, practice seems to increase the correlations slightly, especially in the case of "rate" (Table IX). The situation with respect to "accuracy" is less certain in its interpretation. By reference to Table X it will be seen that the mean daily scores for accuracy did not increase appreciably. From the first to the second day there was a noticeable increase, 40.12 to 50.83, but subsequently the means are almost constant. This was an unforeseen difficulty and one which has probably affected the results of this test very considerably. The interpretation of this situation is undoubtedly to be had in the fact that, during the course of the practice, the subjects, in the main, failed to learn to solve *any more new types of problems.* A given type of problem was either within the ability of the subject at the outset of the

practice or it never became possible for him to make the solution. The improvement, then, perforce, had to come in the direction of the increased speed of solution of the same types of problems with added practice. As is well known, rate scores are less dependent upon intelligence than are accuracy scores, at least for mathematical abilities.

It should be further pointed out that the rate score includes the factor of accuracy as well as rate proper (see Chapter III), and is therefore the better single measure of performance.

The relative variability of the daily accuracy scores remains practically constant save for minor fluctuations which are not constant in direction. This is equally true of the rate scores where improvement is marked.

That the curves of the individuals, if plotted, would not cross to any considerable extent is shown by the correlations of the first and tenth days' performances. These were $0.782 \pm .033$ for rate and $0.804 \pm .030$ for accuracy.

All of the evidence considered, practice in the solutions of abstract problems has not decreased the amount of correlation of this ability with intelligence, and probably such correlations tend to increase with practice within the time limits of the experiments.

GENERAL CONSIDERATIONS

The individual characteristics of the learning processes involved in the three separate tests employed in the present investigation have been discussed at sufficient length in the preceding pages. There yet remains the task of bringing the specific findings together into generalizations regarding the rôle of intelligence in conditioning learning.

The most striking conclusion which can be drawn from these investigations is that it is impossible to generalize about the form of the learning curve *in toto*. We can speak of learning curves in the plural sense, but there seems to be no reason to assume that there exists any one type of learning curve which has universal validity regardless of the mental functions which may be involved and independent of those differences of innate capacity which we

term intelligence or general mental ability. No doubt there are certain characteristics of learning curves which are common to all, such as rapid initial rise, followed by flattening, and final gradual approach to the physiological limits of improvement. What is meant by this categorical denial of any one generalized form of the curve of learning is that evidence has been here presented which suggests at least two important controlling factors in the rate of learning. These two factors are:

(1) *The type of mental function undergoing exercise, and*

(2) *The mental level of the subjects used.*

These issues must also be stated in terms of the problem of the increase or decrease of individual differences under practice. Reference has already been made to the fact that the whole matter of individual differences in learning capacities has never been subject to agreement. Specifically, we can refer again to the findings of Binet that individual differences tended to be effaced with practice, at least for such simple tasks as cancellation tests of various types. Spearman and Krüger have already been quoted to the contrary in their refiguring of Oehrn's data on continued adding. Burt's results were inconclusive in the main and cannot be held to support either contention. On the other hand, Thorndike has adhered firmly to the conviction that those differences present in performance at the outset due to man's original nature are markedly increased by continued exercise. Thorndike would corroborate the work of Spearman and Krüger and oppose the position taken by Binet if these findings are accepted at face value without further analysis.

In the light of certain new data which have been presented here, however, it is perhaps possible to reconcile these divergent opinions in considerable measure. These opposed conclusions have probably not taken into account sufficiently the first of the above mentioned factors which have been emphasized in the present study, viz., the effect of the type of mental function undergoing exercise. Binet's tests were almost entirely of the cancellation type, and it is now well established that such tests involve relatively simple mental processes. The work of Thorn-

dike and his students was almost entirely concerned with relatively complex functions such as mental multiplication and adding. The same holds true for the study of Spearman and Krüger. Moreover, considerable new evidence has been presented in the present discussion which serves to show that both conclusions are probably correctly drawn for the materials used, and that the real error involved is merely that of overgeneralization from a too restricted sampling of mental processes concerned in the learning.

In order to examine this general hypothesis more specifically and from several possible angles, the general issue can be divided into at least three more sharply defined questions. These are:

(1) The question of the changes in the relative variabilities of the entire practicing group, *i.e.*, tendencies for the ratio of standard deviation to the mean to change from day to day during the practice. Such ratios are found in the Pearson Coefficient of Variation.

(2) The problem of the increase or decrease of the individual differences in performances in terms of absolute increments when the practicing group is subdivided into a "superior," an "average," and an "inferior" group upon the basis of mental age.

(3) The question of the fate of correlations between mental age and performance during continued practice.

The changes in the relative variability of the entire practicing group have been indicated by the Pearson Coefficients of Variation as given in Tables VI, VIII, and X. In the case of card sorting the relative variability of the group became rapidly less in the first five trials, being 25.93 for trial 1 and falling to 18.96 for the fifth or last trial of the first day of practice. However, if we consider the average daily scores, there is no constant tendency toward change in either direction. The group as a whole represents a condition of no change in relative variability under practice. The rather rapid change within the first five trials is a very transitory phenomenon, and it must be pointed out that the single trials are much less reliable than the daily averages, and hence are

not strictly comparable to such averages. The chief significance of this stability of the Pearson Coefficients of Variation is probably that they offer evidence that changes in the correlations of mental age and performance are little affected by changes in relative variability from day to day.

In the case of the code substitution test, the situation with respect to relative variabilities is slightly different. In both kinds of substitution, there is a marked decrease in the variability from the first to the second day, and little change subsequent to the second day. This change amounts to a difference of from 53.65 to 39.81 for code to English, and from 50.49 to 35.61 for English to code. Again, variability in the relative sense here implied cannot be a factor in change in the magnitudes of the correlations of intelligence and performance.

Finally, in the case of the abstract relations test there is even less change in the magnitudes of the Coefficients of Variation. For reasons already stated, the rate scores are the better single measure. Although the coefficient for the first day is slightly greater than on any other day, it is almost equalled on the fourth and fifth days, and the differences are probably not significant. In the case of accuracy there is absolutely no reason to believe that systematic tendencies toward change are present in either direction.

Taking all three tests together, the relative variabilities are surprisingly constant over the period of ten days of practice. In the first two tests it would appear that there exists a relatively short period in which the members of the practicing groups tend to arrange themselves in a somewhat more homogeneous manner in their individual performances, but that this is a very transitory phenomenon.

In order to discuss the second of the problems to be considered, viz., that of the increase or decrease of the differences between the three subgroups classified as "superior," "average," and "inferior," certain new data must be introduced. These three groups were formed by breaking the total group into the highest, middle, and lowest thirds taken in the order of their mental ages.

For the three tests the divisions were as follows. The numbers correspond to those assigned to the subjects in the tables of the raw data.

	Card sorting	Code	Abstr. Rel.
"Superior" group........	Nos. 1–17	Nos. 1–22	Nos. 1–21
"Average" group.........	" 18–34	" 23–44	" 22–42
"Inferior" group..........	" 35–52	" 45–66	" 43–63

Tables XIII to XV and Graphs I to V show the mean scores of each of these three groups during each day of the practice.

In the case of card sorting, there is a very marked tendency for the groups to draw together, *i.e.,* to lose the differences present at the outset. This is most rapid during the first five trials. By the second day the three groups have lost their separate identities. In the case of the abstract relations test we find the reverse situation with respect to the absolute gains. In terms of absolute increments the three groups are constantly increasing those differences present in the initial stages of the learning. The curves diverge like a fan. These two tests tend, then, to show opposed tendencies toward convergence or divergence. In the substitution test we find the curves to be rather irregular. However, if the same were smoothed until they approximate best-fitting lines, the

TABLE XIII

Average Time Scores in Seconds for the Superior, Average, and Inferior Groups in Card Sorting

Trials of	Superior		Average		Inferior	
First Day	Sec.	N	Sec.	N	Sec.	N
1	261.4	17	280.8	17	336.7	18
2	176.1	17	189.3	17	218.6	18
3	153.1	17	161.4	17	169.4	18
4	139.6	17	150.6	17	155.7	18
5	134.3	17	140.9	17	146.9	18
Days of Practice						
1	173.1	17	182.9	17	201.6	18
2	111.8	17	109.8	17	117.8	18
3	101.8	17	108.9	17	103.9	18
4	95.7	17	99.0	17	101.2	18
5	93.2	17	102.6	17	97.7	18
6	94.2	16	101.0	16	95.2	17
7	90.3	16	99.6	16	90.9	17
8	89.0	16	98.5	16	87.6	17
9	89.8	17	97.1	16	90.8	16
10	88.8	17	95.9	16	86.7	16

separate curves of the three groups of subjects would be roughly parallel.

The whole range of possibilities in the forms of the learning curves for groups classified upon the basis of mental age can be represented hypothetically by Diagram I which is given here. Type I represents convergence of the curves or decrease of differences in gross improvement under practice. The card-sorting experiment illustrates Type I. Type III presents the situation found to obtain for the abstract relations test. Type II is not exactly represented in our findings, but is theoretically possible, and there is some evidence that the code test approaches this type to some extent.

The third of our three problems is concerned with the fate of correlations between performance and intelligence. It bears certain similarities to each of the other issues which have just been discussed. It has been repeatedly pointed out that changes in the relative variabilities of the practicing group may influence the

TABLE XIV

Average Daily Scores in Code Substitution for the Superior, Average, and Inferior Groups of Subjects

Day	Superior		Average		Inferior	
Code to English	Lett.	N	Lett.	N	Lett.	N
1	134.0	22	102.3	22	59.8	22
2	193.4	22	167.5	22	110.7	22
3	244.3	22	203.7	22	141.9	22
4	254.6	22	206.0	22	162.0	22
5	270.2	22	212.2	22	184.6	22
6	291.0	22	222.4	22	195.5	22
7	306.4	22	237.8	22	195.6	22
8	329.5	22	246.2	22	189.6	22
9	336.8	22	250.7	19	213.0	20
10	346.8	21	256.9	17	222.4	20
English to Code						
1	149.0	22	114.9	22	67.4	22
2	187.5	22	163.0	22	111.3	22
3	215.6	22	183.9	22	133.1	22
4	226.7	22	194.9	22	156.8	22
5	235.4	22	189.3	22	167.3	22
6	248.4	22	200.3	22	174.8	22
7	274.0	22	215.0	22	172.9	22
8	281.5	22	224.8	22	173.9	22
9	301.6	22	225.6	19	185.5	20
10	318.7	21	232.6	17	206.8	20

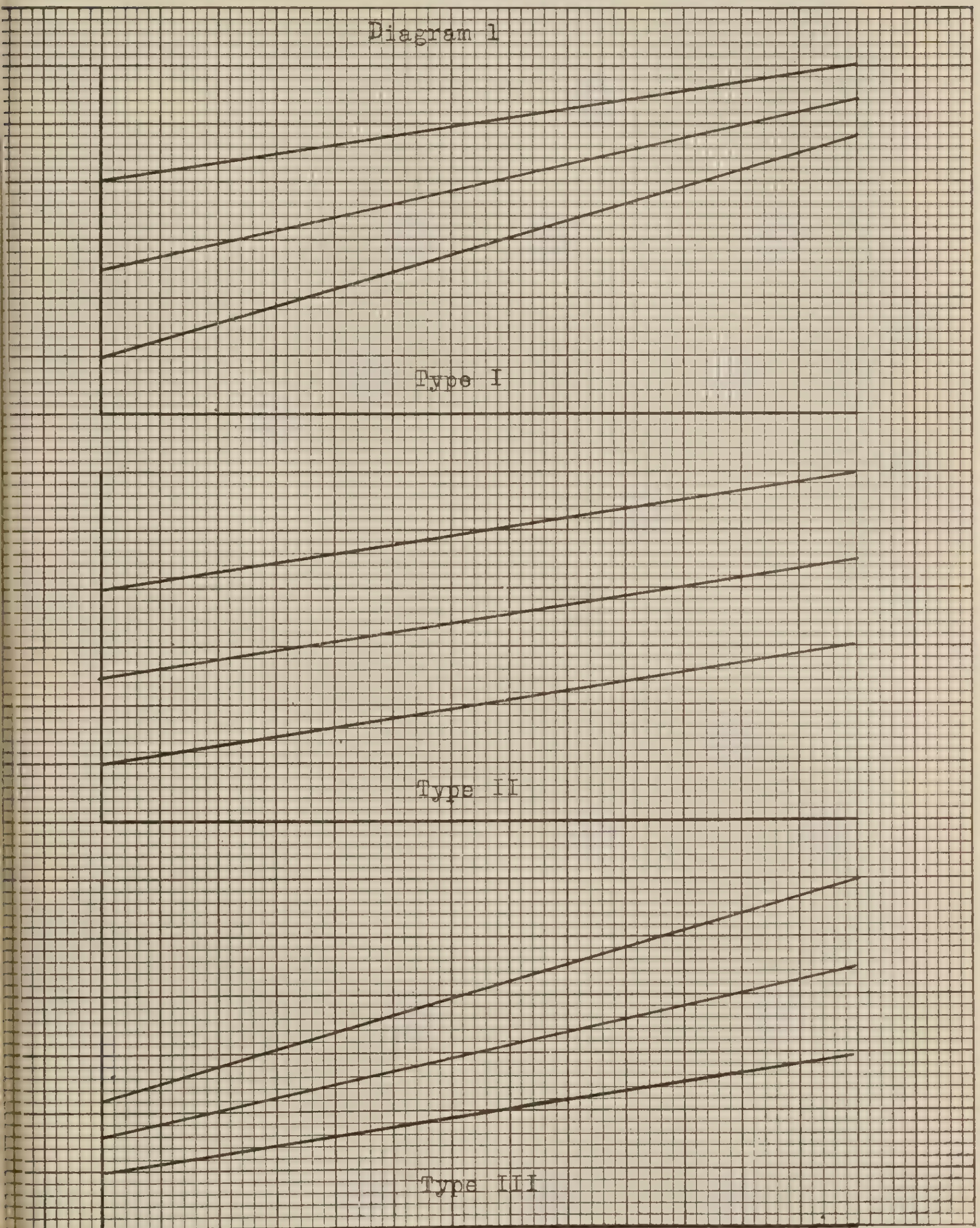

Fig. 3.

magnitude of such correlations because of increase or reduction of the range of talent involved. However, the fact of relative constancy of such relations in all of the tests, at least after the brief initial period, eliminates this factor in the main.

It is, of course, entirely possible, theoretically, that a practicing group can maintain the same relative variability as measured by the ratio of standard deviation to the mean, and yet *the individuals composing the group will be found to approach more and more*

TABLE XV

Average Daily Rate Scores in the Abstract Relations Test for the Superior, Average, and Inferior Groups of Subjects

Day of	Superior		Average		Inferior	
Practice	Prob.	N	Prob.	N	Prob.	N
1	33.9	21	16.4	21	8.8	21
2	63.8	21	27.6	21	14.2	21
3	68.4	21	35.5	21	16.6	21
4	75.4	21	34.5	21	20.9	21
5	88.4	21	41.8	21	22.1	21
6	92.7	21	42.8	21	27.2	21
7	101.7	21	46.6	21	25.3	21
8	110.6	21	52.5	21	31.9	21
9	116.8	21	53.6	20	32.3	21
10	118.5	21	57.2	20	30.2	21

closely to the order demanded by their separate degrees of general mental ability. The reverse of this condition is also possible, *i.e.,* the subjects may day by day lose more and more the relation of their performances to their mental abilities. This, in fact, is what would probably tend to happen where some correlation exists between intelligence and performance in the initial stages of the learning but where automatization sets in rapidly with the result that great changes in the mental functions brought into play in the learning take place during the course of the practice. This point has been repeatedly brought out in laboratory studies.

In general we can expect any one of three situations to be found with respect to the fate of correlations between intelligence and learning:

(1) Decrease in the size of the coefficients of correlation from day to day. If the relative variability of the entire practicing group is constant, this means that the individuals tended to

arrange themselves in their performances in the relative orders suggested by their mental ages at the outset, but that, from day to day, these same individuals became rearranged in orders which were quite or almost unrelated to general intelligence. Individual differences would still be found in the same relative degree as before, but intelligence would

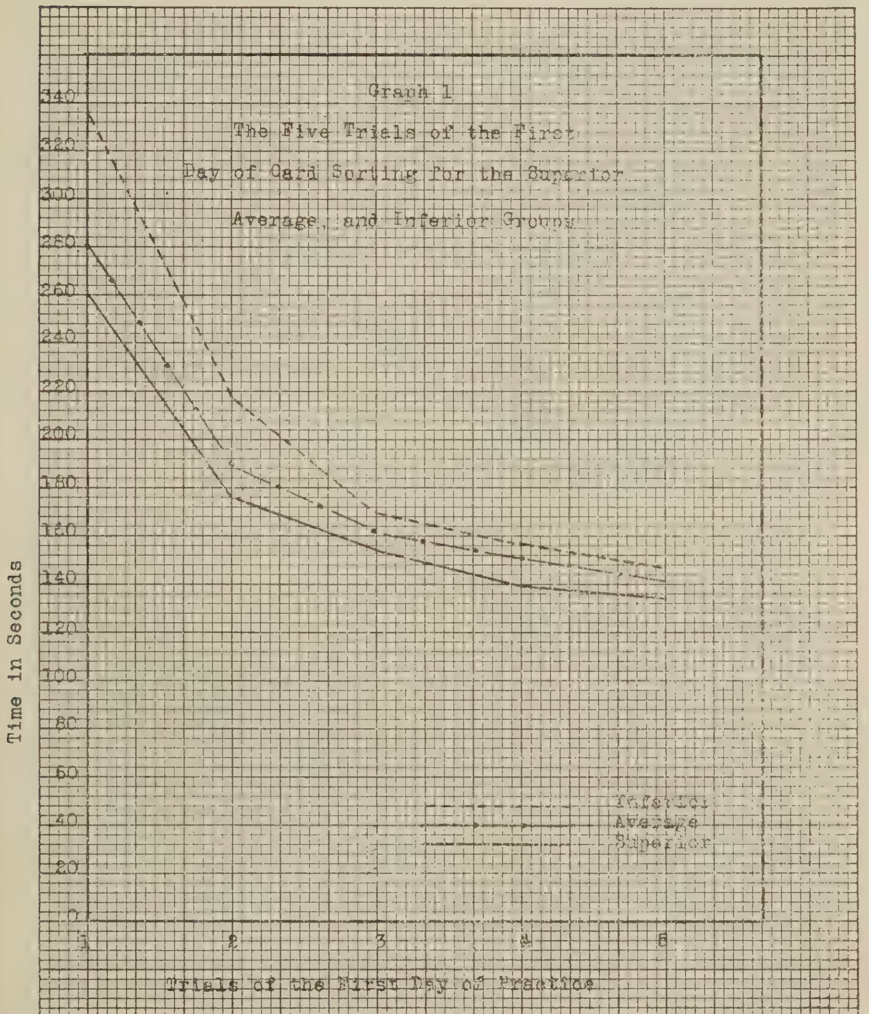

Fig. 4.

not longer be a factor in controlling the positions of the subjects within the group.

(2) The coefficients of correlation might remain constant from day to day. Changes in position of individual members of the group would still be possible, but these changes would probably be relatively small in extent and not constant in direction with reference to the order demanded by the mental

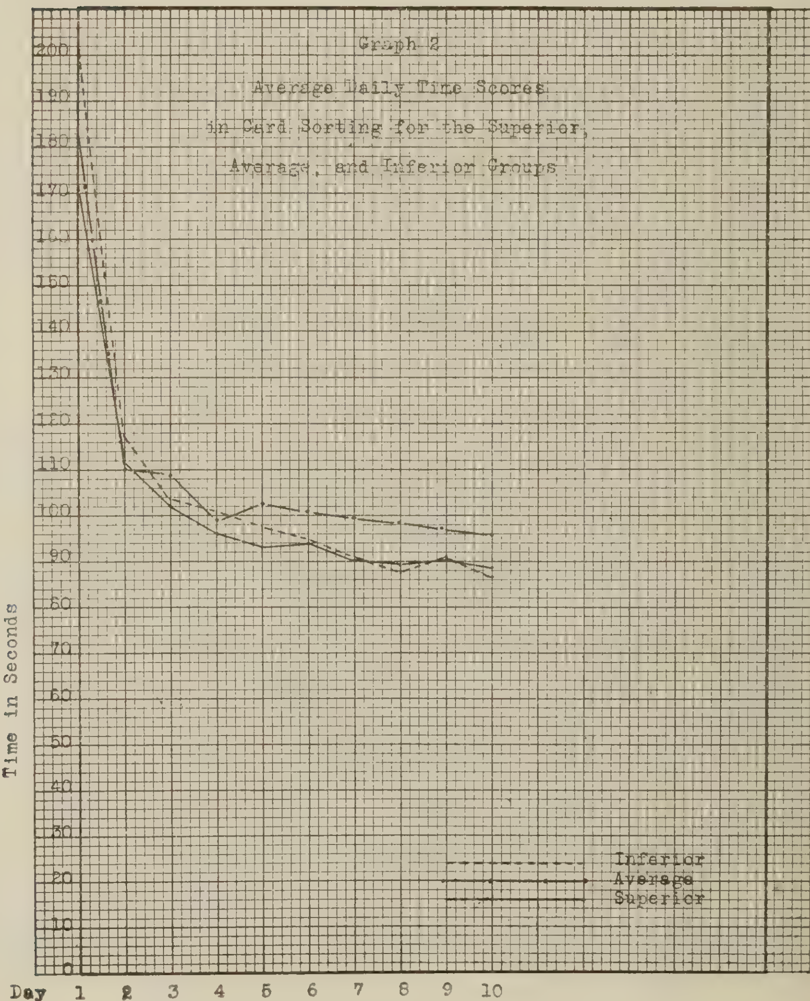

Fig. 5.

ages of the subjects. As before, the relative variability of the group as a whole might remain constant.

(3) The correlations between performance and mental age might be found to be increased with practice. Here again the relative variability of the entire group might not change from day to day, but the individuals composing the group would

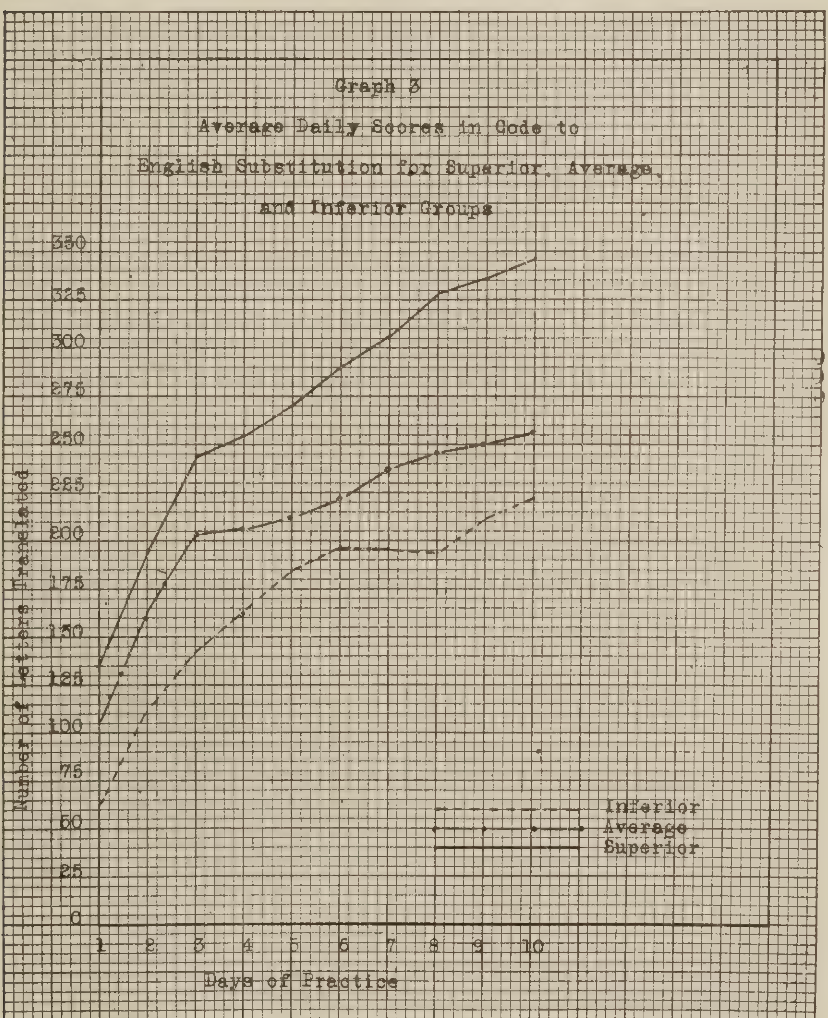

Fig. 6.

be found to be approaching more and more closely the order demanded by their mental ages.

These three possibilities are admittedly hypothetical in part. Each carries with it certain further corollaries. One of these is the question of correlation between initial and final performance.

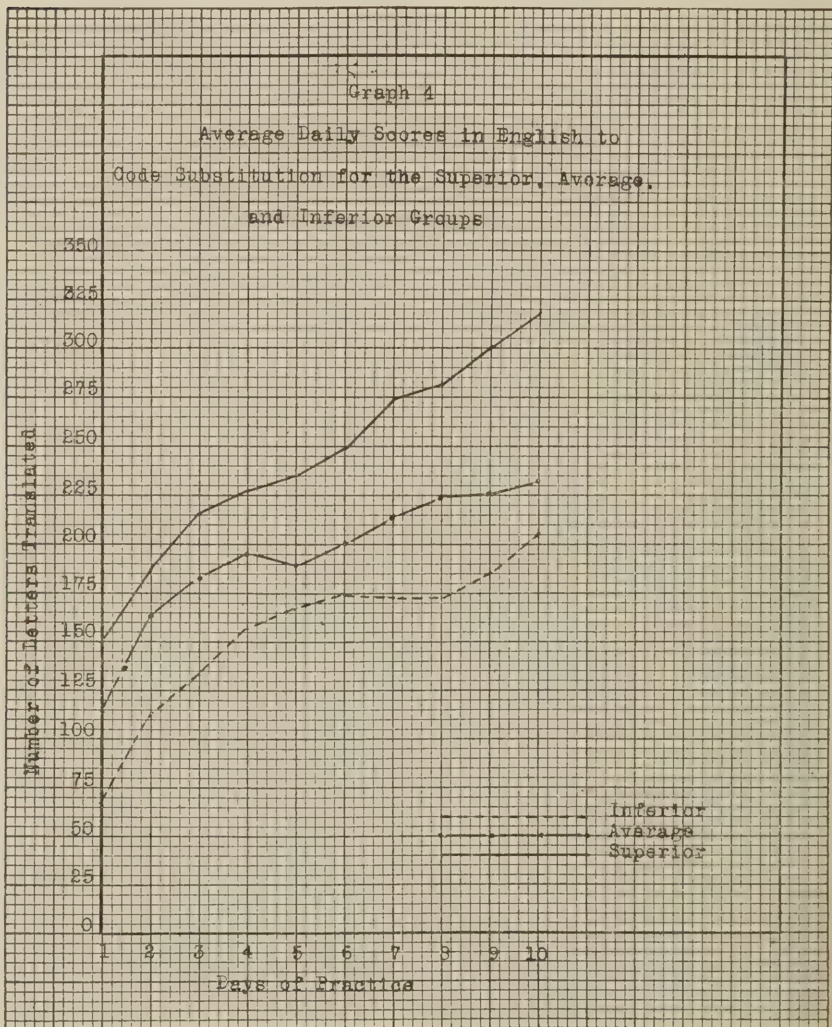

Fig. 7.

The first situation demands a relatively lower correlation between the positions of the subjects at the start and at the finish of practice. The curves of individual subjects are likely to cross and recross toward the end of practice. However, loss of all correlation between mental age and performance is possible with-

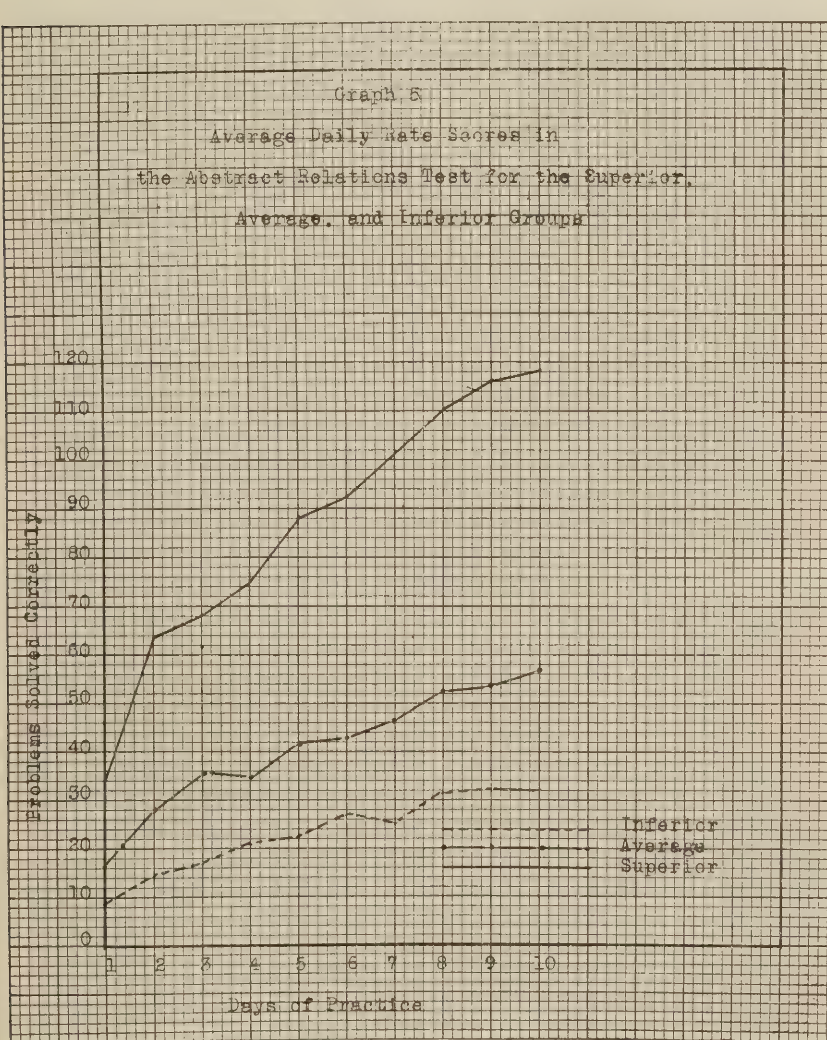

Fig. 8.

out loss of correlation between initial and final positions. This is shown by the card-sorting test where the initial correlation of intelligence and performance rapidly fell to zero although the correlation between the performances of the first and last days was $0.516 \pm .071$. Crossing of the individual curves would probably be slight in the second situation, *i.e.*, where the correlations of performance and intelligence are practically constant. The amount of crossing of individual curves, or what is the same thing, the amount of correlation between initial and final performance, would seem to be controlled by two general factors:

(1) Individual characteristics in the improvability of the particular mental function undergoing exercise. These characteristics would be subject to individual differences in different human beings. This would be individual learning capacity in a mental function *per se*. Such individual differences might conceivably be quite unrelated to the differences existing among the same human beings with respect to intelligence or general capacity for learning. An illustration might be found in the learning to control a baseball by a pitcher. If we assume for the sake of discussion that this is a function entirely unrelated to intelligence, a learning capacity for throwing a baseball accurately might be relatively constant for a given individual and at the same time great individual differences in this power could exist. A moderately high correlation between initial and final ability in this act might be entirely possible. The learning curves of such individuals would not cross to any considerable extent. In fact, this situation is realized to some extent in the card-sorting experiment where the initial performance correlated with the final performance to the extent of $0.516 \pm .071$, but after a very brief initial period the performances on successive days were quite unrelated to intelligence. Individual differences were still as great as before in relative terms.

(2) This first factor would be subject to modification by a second factor, viz., the *general* mental abilities of the subjects. This is really a distinction of the nature of that sometimes assumed by psychologists in the discussions of general versus specific factors in talent. In the case of card sorting, intelligence does not appear to exert any influence upon the positions of the subjects in their relative orders of efficiency

> after the very first few trials. On the other hand, if we consider a complex function like the abstract relations test or even the code test, we find that intelligence is constantly operative in determining the efficiency of the performance. Since intelligence is a constant factor, *i.e.,* does not change in the course of the experimentation to any significant extent, it should act to supplement the specific factors making for constancy of relative position from day to day, and, on the whole, correlations of initial and final positions might be found to be higher than in those cases where only the specific capacities are concerned. The actual correlations for the first and last days' scores in the code test were $0.606 \pm .056$ and $0.779 \pm .035$ for code into English and English into code, respectively. For rate in abstract relations the coefficient was $0.782 \pm .033$ and for accuracy $0.804 \pm .030$.

If both of these factors are varied in opposite directions such as might be the case with a learning situation presenting zero correlation with intelligence at the outset and rising to perfect correlation (1.00) at the end of practice, we would obtain the maximum of crossing and recrossing of the separate curves of different individuals, together with zero correlation between initial and final positions. This example may not be a real one in the sense that it corresponds to any psychologically possible situation. Its introduction is made purely for the purpose of setting the distinction between the specific and general factors in learning into the sharpest possible contrast. Whether such a situation is ever possible is very doubtful.

Comparison of the Present Results with the Work of Certain Other Investigators:

In Chapter II, devoted to the review of the literature related to the present problem, several references were made to conflicting results and interpretations in the experimental studies of learning. In part, these issues have been discussed as our new data were presented. However, several of these questions have not been thrown into orientation with present results. In some cases it has been possible to reconcile differences of opinion; in others, it will be necessary to leave the differences standing in opposition.

The suggestion was previously made that the apparent contradiction between the work of Binet with cancellation tests and that of Spearman and Krüger and of Thorndike on continued adding and other arithmetical functions has resulted from the fact that they were comparing mental functions which are widely different. When this factor is recognized, our present findings can be held to harmonize the two positions up to the point of rejecting the universality of their generalizations. On the other hand, Jones (1917:13) has been quoted to the effect that, in his opinion, the work of Wells, Chapman, and Hollingworth was opposed to the general conclusions of Thorndike. Jones bases his argument primarily on the fact that Wells found that those subjects who gained most in adding over a period of thirty days did not gain the most in cancellation during the same period of practice. Upon the assumption that card sorting is similar to cancellation in its demands upon intelligence and that the solution of problems like our abstract relations problems involves abilities roughly comparable to those of continued adding, the fact of differential gains in Wells' two tests presents no serious difficulty in its interpretation. In our card-sorting test, the dull subjects gained most and the bright ones least. In the abstract relations test, the situation was reversed, the bright ones gained most and the dull ones least. Our results, then, are in harmony with those of Wells. Again, the apparent conflict need not be accepted as real but can rather be attributed to failure to recognize the fact that learning curves vary markedly in form for various types of mental functions.

However, when we consider the differences in the results obtained for correlations between initial and final abilities, it must be admitted that present results are not entirely in harmony with all of the previous work.

Hollingworth (1914:29) reported correlations between initial and final abilities for six tests with thirteen adult subjects, as follows:

Adding	.15
Opposites	—.08
Color naming	.68
Discrimination	.68
Cancellation	.67
Coördination	.52
Tapping	.23

The values obtained for adding and opposites are strikingly different from those reported for the higher mental functions in the present investigation. In opposition to these figures by Hollingworth are Chapman's results (1914:9). The latter found correlations of initial and final performance of .59 for opposites, .96 for addition, .87 for multiplication, .87 for color naming, and .75 to .85 for cancellation.

Neither of these sets of results can be compared directly with those obtained here. In the first place, Hollingworth's subjects continued their practice for 175 trials. As Hollingworth himself has pointed out, the tests underwent great changes in their psychological characteristics during the course of practice. Moreover, the subjects were but thirteen in number in Hollingworth's study and but twenty-two in Chapman's investigation. All were adults. Exactly the same words were used in each practice period although the order of presentation was changed. The conditions, therefore, were favorable for rapid automatization and marked decrease of intellectual demands from day to day. Nevertheless, it must be admitted that there is a considerable disagreement here which cannot be explained at present.

Whitely (1911:54) found correlations between starting point and gain to be equal to about .50 in such functions as cancellation, discrimination of weights, sorting, and the pencil maze. As has been stated, these results are out of harmony with certain of Whitely's other work as well as with the present study. In our card-sorting experiment the subjects making the poorest initial records gained most during practice, a fact which would demand negative correlation. Thorndike has called attention to the fact that Whitely's scores are not very reliable because they were obtained from only nine subjects.

It is unnecessary to enumerate the many other minor disagree-

ments in the literature of the psychology of learning. Enough evidence, it is thought, has been presented in the present paper, to indicate clearly that important generalizations concerning learning capacities must distinguish the separate rôles of general intelligence and the complexity of the mental functions involved, in addition to those specific factors governing changes of efficiency during practice.

CHAPTER VI

Summary and Conclusions

The main conclusions which have been reached in the course of the present experimentation are as follows:

1. There is no reason to assume that there exists *any one type* of learning curve which is independent of:
    - (a) The *type of mental function* which is undergoing practice, and of
    - (b) Differences in the *general mental abilities* of the subjects undergoing the practice.
2. With respect to the second of these two factors, *i.e.*, general mental ability, correlations between performance in learning and measures of general intelligence are subject to any one of three possible fates:
    - (a) Such correlations may decrease from day to day as was the case with card sorting, or
    - (b) Such correlations may remain practically constant from day to day as tended to be the case in code substitution, or
    - (c) Such correlations may increase from day to day. There is some evidence that such a tendency existed in the case of the abstract relations test.
3. When subjects are classified upon the basis of mental age into superior, average, and inferior groups, the separate curves of such subjects may:
    - (a) Converge during practice as was the case in card sorting, or
    - (b) Remain roughly parallel during practice as was practically true of the code substitution test, or
    - (c) Diverge during practice as was shown by the rate scores of the abstract relations test.
4. That the correlation between initial and final performances is controlled by two sorts of factors:
    - (a) Specific capacities *per se,* which are characteristic of that particular mental function, and

(b) The general factor of the degree of alienation between the mental processes involved in the learning and general intelligence. Where the relation is close the correlations of initial and final efficiencies probably tend to be higher than is the case in mental function more distantly related to general mental capacity.

5. That many of the disagreements among students of learning can be harmonized upon the hypothesis that there is no generalized type of learning curve, but that learning curves are specifically conditioned by the type of mental function and by the differences in the general mental capacities of the subjects.

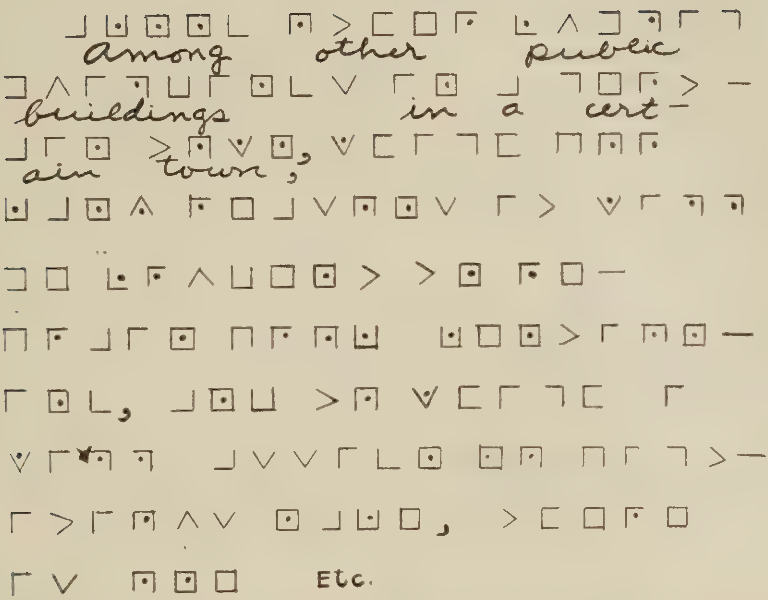

Fig. 9.—A sample of the test materials used for the code to English translation in the substitution experiments. The sample given covers the first few phrases of Chapter I of Oliver Twist. A few of the words have been translated in order to show the method.

A sample of the test materials[1] used for the English to code translation in the substitution experiments. The sample given covers a few phrases from Chapter II of Oliver Twist:

For the next eight or
ten months, Oliver was the
victim of a systematic course
of treachery and
deception. He was
brought up by hand.
The hungry and destitute
situation of the infant
orphan was duly reported
by the workhouse author-
ities to the parish
authorities. The
parish authorities
inquired with dignity of
the workhouse author-
ities, whether there
was no female then

[1] The materials as used were heavily leaded to allow the subjects to write the symbols between the lines.

A sample of the problems used in the abstract mathematical relations test. The first five of the problems have already been answered in order to show the method by which the subjects recorded their answers to the problems.

1	+	B is larger than D
2	+	A plus D is larger than B
3	±	C minus D is larger than B
4	±	C minus A is less than D
5	+	A plus D is less than B plus C
6		A minus B is less than D minus C
7		C minus A is equal to B plus D
8		A plus B plus D is equal to D plus C
9		A plus B plus C is equal to B plus C plus D
10		A minus D is larger than C minus A
11		A is larger than C
12		C plus D is less than B
13		A minus C is larger than B
14		D minus B is equal to A
15		A plus C is larger than D minus B
16		A minus D is less than B plus C
17		C minus A is larger than D minus B
18		A plus C plus D is less than D plus A
19		A plus B plus D is equal to B plus C plus D
20		A minus D is less than B minus D
21		B is less than C
22		A plus C is larger than D
23		A minus D is larger than C
24		D minus C is equal to B
25		A plus D is larger than B minus C
		Etc.

BIBLIOGRAPHY

1. Abelson, A. R.: The Measurement of Mental Ability of Backward Children, *Brit. Jour. of Psych.*, 1911, 4, 268–314.

2. Baldwin, B. T.: The Learning of Delinquent Adolescent Girls as shown by a Substitution Test, *Jour. Educ. Psych.*, 1913, 4, 317–332.

3. Bergstrom, J. A.: An Experimental Study of Some of the Conditions of Mental Activity, *Amer. Jour. of Psych.*, 1894, 6, 247–274.

4. Binet, A.: Attention and Adaptation, *L'annee Psych.*, 1899, 6, 248–404, espec., 362–381 and 402–404.

5. Bonser, F. G.: The Reasoning Ability of Children of the Fourth, Fifth, and Sixth School Grades, *Teach. Coll. Contr. to Educ.*, 1910, No. 37.

6. Brown, Warner: Habit Interference in Sorting Cards, *Univ. of Calif. Pub.*, 1914, 1, No. 4, 269–321.

7. Burt, C.: Experimental Tests of General Intelligence, *Brit. Jour. Psych.*, 1909, 3, 94–177.

8. Calfee, M.: College Freshmen and Four General Intelligence Tests, *Jour. Educ. Psych.*, 1913, 4, 223–231.

9. Chapman, C. C.: Individual Differences in Ability and Improvement and their Correlation, *Teach. Coll. Contr. to Educ.*, 1914, No. 63.

10. Colvin, S. S.: Aspects of the Learning Curve, *Psych. Bull.*, 1915, 12, 67.

11. Coover, J. E., and Angell, F.: General Practice Effects of Special Exercise, *Amer. Jour. Psych.*, 1907, 328–340.

12. Culler, A. J.: Interference and Adaptability, *Arch. of Psych.*, 1912, No. 24.

13. Dallenbach, K.: The Effect of Practice Upon Visual Apprehension in the Feebleminded, *Jour. Educ. Psych.*, 1919, 10, 61–82.

14. Dearborn, W. F., and Brewer, J. M.: Methods and Results of a Class Experiment in Learning, *Jour. Educ. Psych.*, 1918, 9, 63–82.

15. Donovan, M. E., and Thorndike, E. L.: Improvement in a Practice Experiment under School Conditions, *Amer. Jour. Psych.*, 1913, 24, 426–428.

16. Gould, M. C., and Perrin, F. A. C.: A Comparison of the Factors Involved in the Maze Learning of Human Adults and Children, *Jour. Exptl. Psych.*, 1916, 1, 122–154.

17. Gray, C. T.: A Comparison of Two Types of Learning by Means of a Substitution Test, *Jour. Educ. Psych.*, 1918, 9, 143–158.

18. Hahn, H. H., and Thorndike, E. L.: Some Results of Practice in Addition under School Conditions, *Jour. Educ. Psych.*, 1914, 5, 65–84.

19. Hollingworth, H. L.: Correlations of Abilities as Affected by Practice, *Jour. Educ. Psych.*, 1913, 4, 405–414.

20. Hollingworth, H. L.: Individual Differences Before, During, and After Practice, *Psych. Rev.*, 1914, 21, 1–8.

21. Hollingworth, L. S.: *The Psychology of Subnormal Children*, 1920, Macmillan Co., pp. 1–228, espec., Chapter X, How Do the Mentally Defective Learn, 170–189.

22. Johnson, B.: Practice Effects in a Target Test—A Comparative Study of Groups Varying in Intelligence, *Psych. Rev.*, 1919, 26, 300–316.

23. JONES, E. S.: *The Influence of Age and Experience on Correlations Concerned with Mental Tests,* 1917, Warwick and York, pp. 1–89.

24. KIRBY, T. J.: Practice in the Case of School Children, *Teach. Coll. Contr. to Educ.,* 1913, No. 58.

25. KLINE, L. W., and OWENS, W. A.: A Preliminary Report of a Study in the Learning Process Involving Feeling Tone, Transference and Interference, *Psych. Rev.,* 1913, 206–244.

26. KUHLMANN, F.: Experimental Studies in Mental Deficiency, *Amer. Jour. Psych.,* 1904, 15, 391–446.

27. LOUGH, J. E.: Plateaus in Simple Learning, *Psych. Bull.,* 1912, 9, 87–88.

28. MURDOCH, K.: Rate of Improvement of the Feebleminded as Shown by Standardized Educational Tests, *Jour. Appl. Psych.,* 1918, 2, 243, 249.

29. MUNN, A. F.: The Curve of Learning, *Arch. of Psych.,* 1909, 2, 36–52.

30. MYERS, G. C.: Some Variabilities and Correlations in Learning, *Amer. Jour. Psych.,* 1918, 29, 316–326.

31. ORDAHL, L. E., and G.: Qualitative Differences Between Levels of Intelligence in Feebleminded Children, *Jour. of Psycho-Asthenics,* 1915, No. 2, 3–50.

32. PERRIN, F. A. C.: The Learning Curves of the Analogies and Mirror Reading Tests, *Psych. Rev.,* 1915, 26, 42–62.

33. PETERSON, J.: Tentative Norms in the Rational Learning Test, *Jour. App. Psych.,* 1920, 4, 250–257.

34. PYLE, W. H.: *The Examination of School Children,* 1913, pp. 1–70, espec., 18–22.

35. PYLE, W. H.: Is Individual Learning Capacity Constant for Different Types of Material? *Jour. Educ. Psych.,* 1919, 10, 121–128.

36. PYLE, W. H.: *The Psychology of Learning,* 1920, Warwick and York, pp. 1–

37. RUGER, H. A. The Psychology of Efficiency, *Arch. of Psych.,* 1910, No. 15, pp. 1–88.

38. SIMPSON, B. R.: Correlations of Mental Abilities, *Teach. Coll. Contr. to Educ.,* 1912, No. 53.

39. SPEARMAN, C., and KRUEGER, F.: Die Korrelation zwischen Verschiedenen Geistigen Leistungsfähigkeiten, *Zeitschrift für Psych.,* 1907, 44, pp. 50–114.

40. SQUIRE, C. R.: Graded Mental Tests, *Jour. Educ. Psych.,* 1912, 3, 432–4.

41. STARCH, D.: Transfer of Training in Arithmetical Operations, *Jour. Educ. Psych.,* 1911, 2, 306–310.

42. STARCH, D.: Periods of Work in Learning, *Jour. Educ. Psych.,* 1912, 3, 209–213.

43. STRICKLAND, G. I.: The Influence of Practice on the Correlations in Ability, *Jour. Educ. Psych.,* 1918, 9, 393–399.

44. STRONG, E. K.: The Learning Curve as a Diagnostic Measure of General Intelligence, *Psych. Bull.,* 1915, 12, 67–68.

45. TERMAN, L. M.: Genius and Stupidity, a Study of the Intellectual Processes of Seven "Bright" and Seven "Stupid" Boys, *Ped. Sem.,* 1906, 13, 307–373.

46. THORNDIKE, E. L.: The Effect of Practice in the Case of a Purely Intellectual Function, *Amer. Jour. of Psych.,* 1908, 19, 374–384.

47. Thorndike, E. L.: *Educational Psychology,* Vol. III, 1914, espec., pp. 302–307.

48. Thorndike, E. L.: Notes on Practice, Improvability, etc., *Amer. Jour. Psych.,* 1915, 27, 550–553.

49. Thorndike, E. L.: Practice in the Case of Addition, *Amer. Jour. Psych.,* 1910, 21, 483–486.

50. Wallin, J. E. W.: Psychomotor Norms for Diagnosis, *Psych. Rev. Monog. Supp.,* 1916, No. 94.

51. Weidensall, J.: *The Mentality of the Criminal Woman,* 1916, Warwick and York.

52. Wells, F. L.: Relation of Practice to Individual Differences, *Amer. Jour. of Psych.,* 1912, 23, 75–88.

53. Whipple, G. M.: *Manual of Mental and Phyiscal Tests,* 1915, Warwick and York, espec., pp. 133–147.

54. Whitely, M. T.: An Empirical Study of Several Tests for Individual Differences, *Arch. of Psych.,* 1911, 19.

55. Woodrow, H.: Practice and Transference in Normal and Feeble-minded Children, *Jour. Educ. Psych.,* 1917, 8, 85–96 and 151–165.

56. Woodrow, H.: *Brightness and Dullness in Children,* 1919, Lippincott.

57. Woodworth, R. S., and Wells, F. L.: Association Tests, *Psych. Rev. Monogr.,* 1911, No. 57, espec., 53–55.

58. Woolley, H. T., and Fischer, C. R.: Mental and Physical Measurements of Working Children, *Psych. Rev. Monogr.,* 1914, 18, No. 77, espec., 148–184.